

4.1
000-840

**FINAL REPORT – FURTHER DEVELOPMENT, SUPPORT
AND ENHANCEMENT OF CONDUIT**

NASA Ames Project Number NAG 21122

Veronica Moldoveanu and William S. Levine

June 14, 1999

**FINAL REPORT – FURTHER DEVELOPMENT, SUPPORT
AND ENHANCEMENT OF CONDUIT**

NASA Ames Project Number NAG 21122

Veronica Moldoveanu and William S. Levine

June 14, 1999

**FINAL REPORT – FURTHER DEVELOPMENT, SUPPORT
AND ENHANCEMENT OF CONDUIT**

NASA Ames Project Number NAG 21122

Veronica Moldoveanu and William S. Levine

June 14, 1999

.

ACKNOWLEDGMENTS

The final report consists primarily of the edited M.S. thesis of Ms. Veronica Moldoveanu.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
I Introduction	1
II Background	5
1 CONDUIT	8
1.1 Setup Mode in CONDUIT	8
1.1.1 Stability Specifications	12
1.1.2 Gain Margin and Phase Margin Specification	14
1.1.3 Simulation Time-Step	15
1.1.4 Gradient Step-Size	20
1.1.5 Creating new specs with the “Specmaker”	21
1.2 Run Mode in CONDUIT	22

III	Specific Work	26
2	F-14 Problem	27
2.1	Stability and Control Augmentation Systems	27
2.2	Description	29
2.3	Block Diagram	30
2.4	Performance Objectives and Constraints	35
2.5	Setup of the Initialization File	46
2.6	Optimization	47
3	Antenna Problem	55
3.1	Description	55
3.2	Block Diagram	57
3.3	Performance Objectives and Constraints	60
3.4	Setup of the Initialization File	63
3.5	Optimization	63
3.6	Optimization Attempts	72
4	Sensitivity	76
4.1	The Different Forms of Sensitivity	77
4.1.1	Performance sensitivity to plant parameter changes	77
4.1.2	Performance sensitivity to controller parameter changes	77
4.1.3	Parameter sensitivity to small changes in the constraints	83
4.2	Sensitivity Tools for Use When the Gradient of MAX_ACTIVE is zero at the optimal point	84
4.2.1	Insensitivity Ellipsoid	84

4.2.2	Insensitivity	89
4.2.3	Cramér-Rao Bound	91
4.2.4	Eigenvalues and Eigenvectors	93
4.2.5	Positive Semidefiniteness	95
4.2.6	Hessian Formulae	98
4.2.7	Results on Accuracy of Hessian Computation	100
4.2.8	Example	101
4.3	Gradient of MAX_ACTIVE is discontinuous at the optimal point	108
4.4	Parameter sensitivity to small changes in constraints	111

IV Conclusions and Suggestions for Future Work 113

LIST OF TABLES

2.1	Comparison between the optimal points obtained by CONDUIT for the F-14 problem	48
3.1	Comparison between the optimal points obtained by CONDUIT for the antenna problem	66
4.1	Grumman F-14, Hessian's eigenvectors in the optimal point	108

LIST OF FIGURES

1.1	Example of unstable closed-loop system for which MATLAB indicates good stability margins	13
1.2	Gain margin and phase margin for the system in Fig. 1.1 evaluated by the MATLAB function “margin”	13
1.3	Nyquist plot for the system of Fig. 1.1	14
1.4	Nyquist plot for Example 2 as obtained from MATLAB without specifying the frequency range	16
1.5	Nyquist plot for Example 2 on the $[10^{1.1}, 10^3]$ frequency interval	16
1.6	Nyquist plot for Example 2 on the $[10^2, 10^4]$ frequency interval	17
1.7	Bode plot for Example 2	17
2.1	Grumman F-14 pitch-axis control system	31
2.2	Alpha response model	32
2.3	Dryden wind gust model	33
2.4	Comparison between YF-17 criterion and analytic curve of dc-gain*LOS. This is the gain spec. The YF-17 Level 1 region is indicated in black.	41

2.5	Comparison between YF-17 criterion and analytic curve of dc-gain*LOS. This is the phase spec. The YF-17 Level 1 region is indicated in black.	42
2.6	Grumman F-14, nominal starting point	50
2.7	Grumman F-14, nominal point improved by CONDUIT, epsilonF=0	51
2.8	Grumman F-14, nominal point improved by CONDUIT, epsilonF=0.0002	52
2.9	Grumman F-14, trivial starting point	53
2.10	Grumman F-14, trivial point improved by CONDUIT, epsilonF=0	54
3.1	Antenna control system	58
3.2	Antenna problem, nominal starting point	68
3.3	Antenna problem, nominal point improved by CONDUIT, epsilonF=0.0002	69
3.4	Antenna problem, trivial starting point	70
3.5	Antenna problem, trivial point improved by CONDUIT, epsilonF=0.0002	71
3.6	Antenna problem, nominal starting point	73
3.7	Antenna problem, nominal point is driven unstable when stability specs are check only, epsilonF=0.0002	74
4.1	The minimum is Phase 1	78
4.2	The minimum is Phase 1	78
4.3	The minimum is Phase 2	80
4.4	The minimum is Phase 2	80
4.5	The minimum is Phase 2	81
4.6	The minimum is Phase 3	81

4.7	The minimum is Phase 3	82
4.8	The minimum is Phase 3	82
4.9	Two-dimensional insensitivity ellipsoid	88
4.10	Geometric interpretation of insensitivities and Cramér-Rao bounds	90
4.11	Projections	94
4.12	Ellipsoid is degenerate	96
4.13	Ellipsoid is degenerate	97
4.14	Grumman F-14, performance display in the Pcomb chart	102
4.15	Grumman F-14, Cramér-Rao bounds	103
4.16	Grumman F-14, insensitivities	104
4.17	Grumman F-14, correlations	105
4.18	Grumman F-14, Hessian's eigenvalues in the optimal point	107

Part I

Introduction

From the first airplanes steered by handles, wheels, and pedals to today's advanced aircraft, there has been a century of revolutionary inventions, all of them contributing to flight quality. The stability and controllability of aircraft as they appear to a pilot are called flying or handling qualities. Many years after the first airplanes flew, flying qualities were identified and ranked from desirable to unsatisfactory. Later on engineers developed design methods to satisfy these practical criteria.

CONDUIT, which stands for Control Designer's Unified Interface, is a modern software package that provides a methodology for optimization of flight control systems in order to improve the flying qualities.

CONDUIT is dependent on the optimization engine called CONSOL-OPTCAD (C-O)[1]. C-O performs multicriterion parametric optimization. It was developed at the Institute for Systems Research (ISR) at the University of Maryland at College Park. C-O was successfully tested on a variety of control problems before it was tested on the design of a rotorcraft flight control system for the UH-60A helicopter in hover based on the ADOCS controller structure [3]. The optimization-based computational system, C-O, requires a particular control system description as a MATLAB file and possesses the ability to modify the vector of design parameters in an attempt to satisfy performance objectives and constraints specified by the designer, in a C-type file.

After the first optimization attempts on the UH-60A control system, an early interface system, named GIFCORCODE (Graphical Interface for CONSOL-OPTCAD for Rotorcraft Controller Design) [7] was created. This interface allowed the designer to push buttons in order to command the optimization engine, instead of typing keywords in the C-O language. Posing the problem

still required code writing.

After that a big step followed, transforming GIFCORCODE into CONDUIT, [2]. CONDUIT eliminates the code writing, by providing (a) a SIMULINK window to design the block diagram, (b) a large library from which to choose the handling qualities specifications, (c) a “specmaker” facility to create new performance specifications as well as (d) a complete methodology to analyze and test the design. Using CONDUIT any control system problem related to aircraft can be posed just by using the mouse to click on menus and drag objects from libraries. Furthermore, the program provides the designer, after each iteration of C-O, with a very efficient picture of all performance criteria plus supporting plots for each individual criterion, a plot showing the evolution of the design parameters, and a specifications-evolution plot. To benefit from hardware speed and advanced visualization characteristics, CONDUIT was developed on the Silicon Graphics IRIX operating system.

CONDUIT has been used to analyze and design:

- a flight control system for the RASCAL UH-60A helicopter in hover having the ADOCS control structure [3]
- a flight control system for the RASCAL UH-60A in forward flight at 80 knots [4]
- a longitudinal control system for the X-29A Forward-Swept Wing Fly-by-Wire Demonstrator
- a roll channel control system for the XV-15 Tiltrotor aircraft in forward flight mode

- a longitudinal/vertical control system for the Kaman SH-2 rotorcraft in hover
- a lateral/directional control system for the Boeing KC-135 Jet Transport
- a longitudinal control system for the Grumman F-14 aircraft
- an antenna control system

In March 1998 CONDUIT was released to US aircraft and helicopter companies, where it is being used in the design of a number of practical aircraft control systems.

Part II

Background

Computer-Aided Control System Design A reasonable way to design a control system for a complex large system is by using an optimization technique requiring two phases. In the first phase the designer formulates the problem. Such a formulation normally consists of an appropriate system structure, a set of objectives, and reasonable initial values for the adjustable design parameters. During the second phase, an optimization package is used to obtain the best values of the design parameters. The second phase is interactive, allowing the designer to explore the design tradeoffs.

Aircraft and helicopter control systems are complex and designing such systems usually requires balancing competing objectives. Moreover the process of optimizing the performance of the control system for aircraft and helicopters has to be repeated tens (or even hundreds) of times corresponding to the changes in the system that occur at different flight conditions.

CONDUIT performs computer-aided control system design using this optimization technique and it uses C-O[1] as the optimization engine. It is also true that many hypothetical designs may be considered—both for the aircraft and the controller.

Specifications for Helicopters and Aircraft Control systems are present in almost all modern aircraft and helicopters in order to 1) enhance aircraft mission capability, 2) improve handling qualities, and 3) decrease pilot workload.

Control law design can only be performed satisfactorily if a set of design requirements or performance criteria is available. In the case of control systems for piloted aircraft generally applicable quantitative design criteria are difficult to obtain because the ultimate test of handling qualities is the pilot's judge-

ment. A pilot's opinion of the flying qualities of an aircraft is influenced by the ergonomic design of the cockpit controls, the visibility from the cockpit, the weather conditions, the mission requirements, and physical and emotional factors. The variability introduced by all these factors can be reduced by averaging test results over many flights and pilots. A systematic approach to flying qualities evaluation is available through pilot opinion rating scales such as the Cooper-Harper scale [19]. Once a rating scale has been established it is possible to begin correlating the pilot opinion rating with the properties of the aircraft dynamic model, and hence derive some analytical specifications that will guarantee good flying qualities.

The acceptability of flying qualities is quantified in terms of "Levels" that are defined for each specific mission task, using the Cooper-Harper scale. The Cooper-Harper scale consists of pilot ratings from 1 to 10, where 1 indicates "excellent, highly desirable" and 10 means "major deficiencies and lost control during some portion of the required operation". Based on this scale "Levels" are defined as follows: Level 1 indicates pilot ratings between 1 and $3\frac{1}{2}$, Level 2 between $3\frac{1}{2}$ and $6\frac{1}{2}$, and Level 3 between $6\frac{1}{2}$ and $8\frac{1}{2}$. Each Level value represents a minimum condition necessary to meet that "Level" of acceptability. Level 1 indicates that aircraft characteristics are "good, may be some negligible deficiencies" [19], Level 2 shows "moderately objectionable deficiencies" [19], and Level 3 shows "major deficiencies" [19]. An aircraft that exhibits Level 1 flying qualities throughout its flight envelope is fully satisfactory. Some real aircraft exhibit some Level 2 flying qualities. Any Level 3 flying qualities are unsatisfactory.

Chapter 1

CONDUIT

The objective of the software package CONDUIT is to act as a designer's assistant/associate for the analysis, testing, and optimization of flight control systems in order to improve the flying qualities. In aircraft design, the design specifications (and constraints) are often competing and CONDUIT helps the designer to perform the trade-offs among them.

CONDUIT is built on top of the MATLAB/SIMULINK system modeling and analysis environment and the C-O computer-aided, optimization-based parametric design software package. It includes a graphical block diagram editor, a graphical spec editor, a spec library, an editor for the initialization file and multiple layers of supporting analysis plots. This avoids the manipulation of MATLAB "m" files or C-O C-type files. The user only has to select and drag and the system automatically updates the relevant scheme, spec or plot.

CONDUIT has two modes of operation: setup and run.

1.1 Setup Mode in CONDUIT

The setup of a problem in CONDUIT consists of three parts:

1. First the designer has to define (or import) into SIMULINK a simulation of the aircraft and the controller. The real aircraft control system consists of many controller loops, all of them controlling the high-order, nonlinear aircraft dynamics. Usually people do not use the nonlinear dynamics in design because it is too complex and needs special software for simulation. Instead they use linearized dynamics. CONDUIT uses SIMULINK for simulation, so that the block diagram can include nonlinearities in the dynamics. The fact that some nonlinearities, especially saturations and dead-zones, are important to aircraft control system design is one motivation for the CONDUIT design approach. Designs based on a purely linear model of the controlled aircraft that ignore saturation, delays, and dead-zone may be overly optimistic.

There usually exist several models for the same system, corresponding to particular flight conditions (speed: hover, low speed, forward flight or mission tasks: precision tasks, aggressive task, etc.). For a thorough examination of the problem, the control systems of all these models should be investigated and optimized.

The control law model must include the design parameters such as gains and time constants, as well as the inputs and outputs of the system. The inputs are used to apply test input signals. The outputs are used to evaluate system responses.

2. The next step for the designer is to define the set of design specifications. To help the designer, there are five graphical libraries comprising over 50 specifications in CONDUIT. These are organized into the following sub-

libraries:

- Generic Specifications
- Fixed-Wing Longitudinal
- Fixed-Wing Lateral/Directional
- Rotor Hover Low Speed
- STOVL (Short Take-Off / Vertical Landing)

Most of the specifications belong to some aeronautical design standards: rotorcraft specifications belong to ADS-33D [19] and aircraft specifications belong to MIL-STD-1797 [20] and Mil STD 9490.

The designer only has to select the appropriate design specification from the available libraries, drag them into the HQ Window (Handling Qualities Window), and configure the simulation appropriately for each specification using the HQ Editor (Handling Qualities Editor).

All the specs are created using the same color map: the Level 1 region is blue; the Level 2 region is magenta; the Level 3 region is red. The splines that define the boundaries between levels can be graphically altered to design new specifications.

To “wire” a specification to the SIMULINK simulation model the designer has to use the graphical “HQ editor” . Here, the user declares each specification to belong to one of the following five classes: 1) hard constraint, 2) soft constraint, 3) objective, 4) summed objective, or 5) check only. Those specifications that are declared to be “summed objective” are treated as soft constraints until the design reaches Level 1. Then, all the summed

objective specs are added together and treated as a single objective. At this stage, an improvement in any component of the summed objective improves the design's performance. The selection of the specification class defines the problem for the C-O optimization process. In aircraft design, examples of choosing the specifications are the following: stability specifications are hard constraints; most of the other specifications are soft constraints (bandwidth and time delay (frequency-domain point spec), quickness (time-domain point spec), attitude hold (time-domain envelope spec), interaxis coupling (time-domain point spec and envelope spec)); crossover frequency, actuator energy, and actuator saturation are usually performance objectives.

The input and output port connections for each specification are indicated in an information box in the spec editor.

In case the designer needs to add a specification that is missing from the library he can create it and add it to the library with the "spec maker" facility.

3. The designer must also set up a small initialization file to define problem-dependent constants such as the simulation time-step, the performance precision, denoted by ϵ_F , and the test input signals.

In the following paragraphs some details regarding the problem setup will be presented. "Stability specification" and "Gain margin and phase margin specifications" refer to the specs that must be included in the HQ Window in order to insure a stable system. "Simulation time-step" and "Gradient step-size" refer to some options that need to be specified in the initialization file.

1.1.1 Stability Specifications

Stability criteria are the most important specifications for aircraft control system design and therefore are the very first selected in the performance chart. The “Generic Specifications” section of the CONDUIT spec library contains two specs designated to evaluate the system stability: EigLcG1(eigenvalue test) and StbMgG1(gain and phase margin test). Usually these two specs should be used together. Using only the eigenvalue test, it is not possible to determine the classical stability margins that define robust stability. However, the eigenvalues are a definitive test of closed-loop stability. In contrast, the classical gain and phase margins are insufficient to determine stability. In part this is because they apply only to specific loops, rather than the complete closed-loop system. This is also due to errors in MATLAB. MATLAB computes gain and phase margin based on a Bode diagram without first checking the stability. This makes it possible that unstable systems appear to have good gain and phase margin, as in the following example.

Example 1: Take the system in Figure 1.1. Checking in MATLAB:

```
>> [g, p] = bode(num, den, logspace(-1, 5, 1000))  
>> margin(g, p, logspace(-1, 5, 1000))
```

The MATLAB function “margin” shows gm=14.18 and pm=45.05, as shown in Figure 1.2, which are good stability margins. The poles of the closed-loop system are:

-10.5415, -0.7551+2.2655i, -0.7551-2.2655i and 0.0516, so the system is unstable.

The Nyquist plot constructed for this system, with the contour modified to take a small detour around the pole, at $s=0$, is shown in Figure 1.3. It is apparent that the polar plot of the open-loop system encircles the $-1+j0$ point once as ω

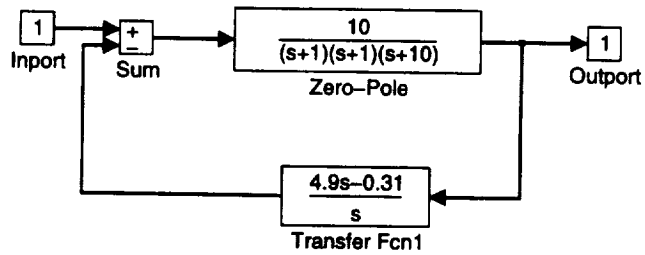


Figure 1.1: Example of unstable closed-loop system for which MATLAB indicates good stability margins

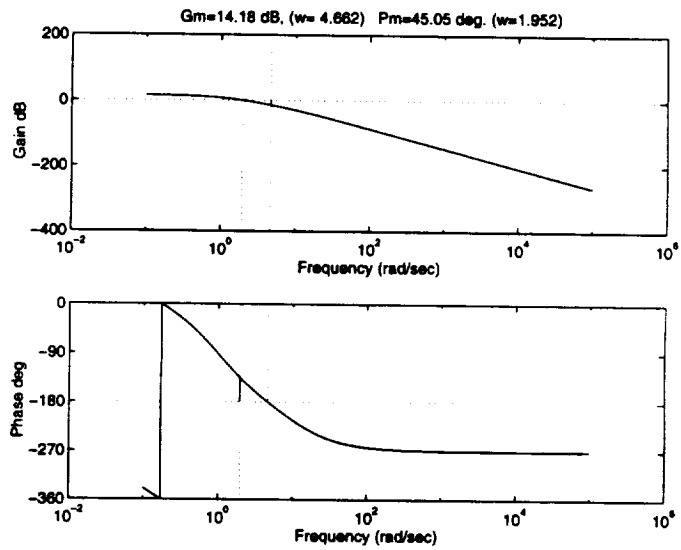


Figure 1.2: Gain margin and phase margin for the system in Fig. 1.1 evaluated by the MATLAB function “margin”

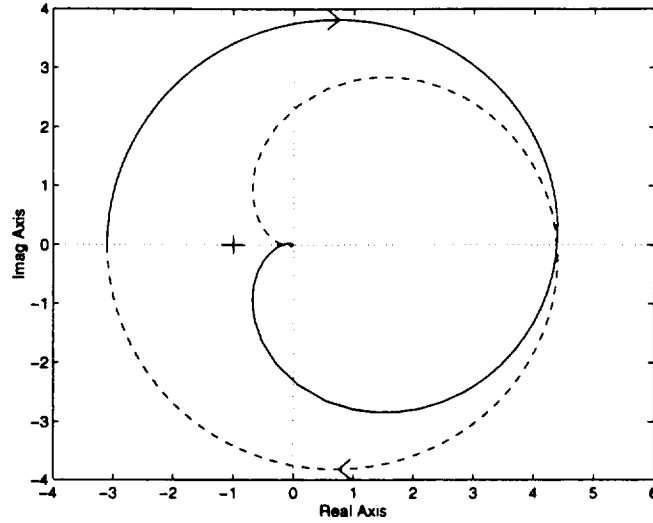


Figure 1.3: Nyquist plot for the system of Fig. 1.1

varies from $-\infty$ to $+\infty$ and thus this closed-loop system is unstable.

The solution to eliminate the instability is to use the eigenvalue test. It is recommended to use EigLcG1 and StbMgG1 both as hard constraints, to be evaluated and optimized in Phase 1.

1.1.2 Gain Margin and Phase Margin Specification

There are some stable systems that have both a gain margin increase and a gain reduction margin, i.e. both the gain increase and gain decrease must be limited. CONDUIT allows one value for the gain margin, so in this case it will take the smallest absolute value of the two of them as the gain margin. This means that the system can increase or decrease by that gain margin and still remain stable.

Example 2:

The antenna problem, part III, chapter 3, has two gain margins for the following

parameters, called the nominal point:

$$\begin{aligned}
 dpp_1 &= 1.08000e + 02 & dpp_7 &= 5.00000e + 02 \\
 dpp_2 &= 5.40300e + 01 & dpp_8 &= 5.00132e + 03 \\
 dpp_3 &= 2.62000e + 03 & dpp_9 &= 7.54000e + 02 \\
 dpp_4 &= 1.57100e + 03 & dpp_10 &= 2.00000e - 01 \\
 dpp_5 &= 1.41370e + 00 & dpp_11 &= 6.67500e - 01 \\
 dpp_6 &= 4.00000e + 00 & &
 \end{aligned}$$

This system is stable, having all poles in the left half plane. Figure 1.4, Figure 1.5, and Figure 1.6 show that the system is conditionally stable. The polar plot of the open-loop transfer function cuts the real axis at -0.4636 and at -48.4172 . This shows that the

gain can increase by $1/0.4638$ ($20 \cdot \log_{10}(1/0.4638) \text{dB} = 6.67 \text{dB}$) and the gain can reduce by 48.4172 ($20 \cdot \log_{10}(1/48.4172) \text{dB} = -33.7 \text{dB}$)

before instability occurs and defines both a gain increase margin and a gain reduction margin. Figure 1.7 shows the gain reduction margin.

For this case CONDUIT will take $\max(\text{abs}(6.67), \text{abs}(-33.7)) \text{dB} = 6.67 \text{dB}$ as the gain margin. CONDUIT uses its own “margin” function that fixes the problems with MATLAB’s “margin” function. MATLAB’s “margin” function doesn’t show all the gain margins and phase margins.

1.1.3 Simulation Time-Step

The real aircraft control system consists of many digital controller loops, with the inner controller loops faster than the outer controller loops, all of them controlling the high order, nonlinear aircraft dynamics. Designers prefer to work with continuous systems because they are easily understandable. Simulations,

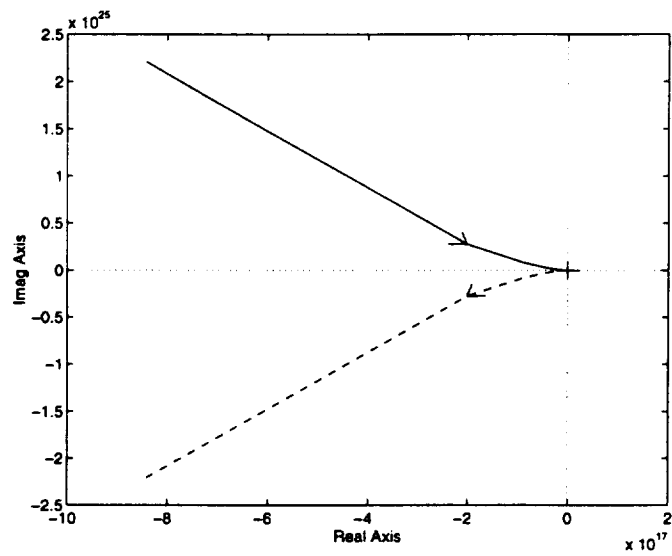


Figure 1.4: Nyquist plot for Example 2 as obtained from MATLAB without specifying the frequency range

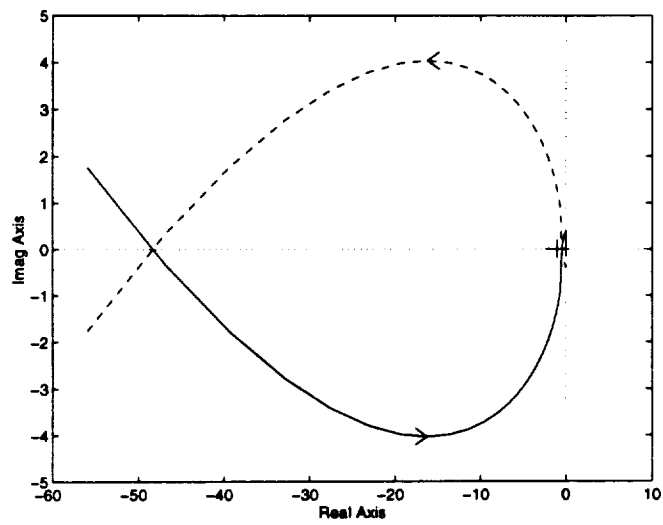


Figure 1.5: Nyquist plot for Example 2 on the $[10^{1.1}, 10^3]$ frequency interval

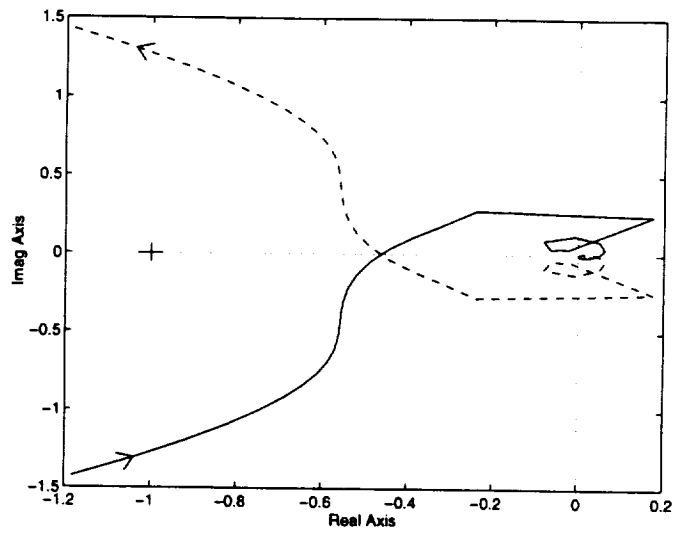


Figure 1.6: Nyquist plot for Example 2 on the $[10^2, 10^4]$ frequency interval

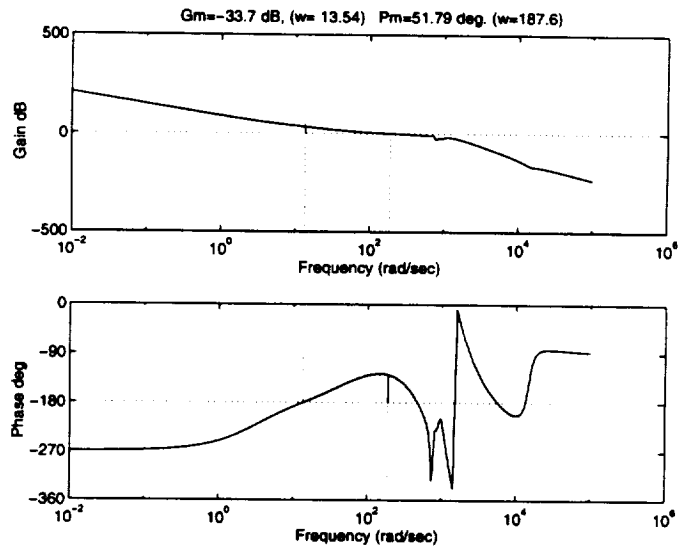


Figure 1.7: Bode plot for Example 2

such as those produced by SIMULINK, run faster using discrete systems. Then there is a choice to use in simulation:

1. Use a continuous model and continuous controller

This has the advantage of fidelity to the real aircraft - aircraft fly in continuous time. Also, most designers have more experience with continuous-time systems and better intuition. Lastly, the specifications are based on continuous-time behavior.

However, the simulation of continuous-time systems is relatively slow. Moreover, the actual implementation of the controllers is in discrete time.

2. Use a continuous model and digital controllers, each with its own sample rate

This has the advantage of very accurate results because of the use of realistic digital controllers, but slow simulation.

3. Use a continuous model and continuous controllers, discretized at a single sample rate

This has the advantage of choice 1 as well as a relatively fast simulation. The disadvantage is that the solution obtained is less accurate. However, it is easy to obtain sufficient accuracy by appropriate choice of the step size.

CONDUIT uses the third option. The designer has to provide a continuous model and continuous controllers for the block diagram. The designer must also choose a sampling interval T , for the simulation. The value of T is placed in the initialization file. Some general rules of choosing T are:

1. An absolute lower bound for the sample rate ($1/T$) is stated by the Shannon sampling theorem [12]. If it is desired that the closed-loop system have a certain bandwidth, say ω_b , then the sample rate must be at least twice this required closed-loop bandwidth.
2. If white noise disturbances are the dominant source of error in the system, then the sample rate should be faster. Theoretically, all real signals have spectral content at all frequencies so there is always information lost due to sampling. By [12], testing the RMS (root mean squared) value of the steady-state value of the covariance of one output, $\sqrt{E[y^2]}$, for the same design and design parameters, but different sample rates, shows that for good random disturbance attenuation a sampling interval of $T = \frac{\pi}{10\omega_b}$ would be a good choice. The relative errors grow quickly when sampling slower than this multiple, whereas very little can be gained by sampling faster. Thus,

$$\frac{\pi}{10\omega_b} < T < \frac{\pi}{\omega_b} \quad (1.1)$$

CONDUIT proposes that the sample rate $\frac{1}{T}$ be 20 times bigger than the fastest mode of the closed-loop system, determined from the eigenvalues of the closed-loop system. Usually the fastest mode for aircraft lies between 2Hz and 10Hz, so the sampling interval, T , is between .005 secs and .25 secs.

The optimal controllers computed by CONDUIT are based on a discretized version of the aircraft and its controller. The discretization uses sampling interval T . the controller actually used by CONDUIT is shown in the dtest.m block diagram in CONDUIT.

The designer should be aware that for different values of T in the initialization file, CONDUIT can provide slightly different controllers.

1.1.4 Gradient Step-Size

C-O evaluates the gradients of the optimization specs at every iteration during the optimization process. The increment used to compute gradients in theoretical problems for which the data is perfectly accurate, is

$$ppdelta = \max(rsteps * \max(1.e0, fabs(x[i])), pdelta1) \quad (1.2)$$

where $x[i]$ is the parameter perturbed in computing the gradient, $fabs$ is the absolute value, $rsteps$ is the square root of machine precision, $pdelta1$ is an arbitrary constant step, usually 0. This formula shows that the gradient step-size is usually $rsteps * fabs(x[i])$. It is $rsteps$ if $x[i]$ is smaller than 1 and it is $pdelta1$ if the user chooses a bigger step.

For real problems, such as those optimized by CONDUIT, the increment used is

$$\sqrt{\epsilonpsilon F} * \max(1.e0, fabs(x[i])) \quad (1.3)$$

where $\epsilonpsilon F$ is chosen by the designer and represents the precision of the optimization specs (C-O allows different $\epsilonpsilon F$ between specs). This way the gradient step-size is usually $\sqrt{\epsilonpsilon F} * fabs(x[i])$ and it is $\sqrt{\epsilonpsilon F}$ if $x[i]$ is smaller than 1. The advantage of using $\sqrt{\epsilonpsilon F} * fabs(x[i])$ instead of $pdelta1$ is that this bigger step is proportional to the parameter, not a constant value. Moreover it can be set up from the initialization file. The current implementation of CONDUIT, to simplify the designer's job, allows only one $\epsilonpsilon F$ for all the specs. The designer has to set the value in the initialization file. The default value is 0.0002 and was chosen following a set of tests on aircraft optimization

problems [24]. Note that this is substantially larger than the increment used in theoretical problem.

1.1.5 Creating new specs with the “Specmaker”

The “Specmaker” facility gives flexibility to the designer to create and use any performance criterion he wants in CONDUIT. There are 5 types of specs that can be added to CONDUIT [25].

- time point
- frequency point
- time line
- frequency line
- LOES (lower-order equivalent system)

For time point and frequency point specs the performance criterion must consist of 2 lines that separate the 3 levels of performance. For time line and frequency line specs the performance criterion must consist of 2 envelopes (closed curves) that encircle the Level 1 and Level 2 regions, respectively. Outside both envelopes is Level 3. Examples are given in the following chapters.

Specs based on lower order equivalent systems (LOES specs) are common in aircraft control. The basic idea is as follows. Given a time or frequency response plot obtained either experimentally or from a detailed high-order simulation. Approximate this given response by the response of a lower-order and much simpler system. The lower order system is characterized by a relatively small number of parameters. These parameters are chosen to minimize some precise

measure of the difference between the lower order model and the real system. The actual specification is given in terms of the lower order model.

1.2 Run Mode in CONDUIT

Once the problem is set up the second phase, design tuning, follows. During this phase the designer cannot change the simulation model, performance specifications, or initialization file. But he can change the current design parameters, evaluate the specifications for every set of design parameters, and use the analysis functions of the system. And CONDUIT will run C-O to optimize the design parameters.

A distance algorithm in CONDUIT translates the location of the design point (in case there is a simple specification) or the location of the worst design point (in case there is a functional specification) on each of the graphical specification criteria to a numerical rating. All of C-O's "good-bad" factors are scaled - the scaled good values become 1 and the scaled bad values become 2. A rating of "1" for the scaled design point indicates that it lies on the Level 1/2 border. A rating of "2" indicates that the design point lies on the Level 2/3 border. These numerical ratings are used by C-O to tune the design and they appear in the Pcomb (Constraint Table). CONDUIT feeds the specifications in the form of the "problem description file" to C-O.

C-O implements multicriterion parametric optimization. It uses FSQP (Feasible Sequential Quadratic Programming) [1] to solve the following general optimization problem:

$$\min_x \text{obj}_i^s(x) \quad \forall i \quad (1.4)$$

$$\begin{aligned}
\text{subject to :} \quad & \text{soft}_j^s(x) \leq 0 \quad \forall j \\
& \text{hard}_k^s(x) \leq 0 \quad \forall k \\
& \text{hard_bound}_l(x) \leq 0 \quad \forall l
\end{aligned}$$

where obj_i^s , soft_j^s , hard_k^s are the scaled values of objectives, soft constraints, hard constraints (except hard bounds), respectively and where hard_bound_l represents hard bounds, an important special case of hard constraints.

- An objective is a specification of a quantity that should be optimized (minimized or maximized). Typically, multiple competing objectives are present.
- A hard constraint is a specification of a quantity that must achieve a specified threshold. A design for which a hard constraint is not satisfied cannot be satisfactory.
- A soft constraint is a specification of a quantity that should achieve , or at least approach, a specified threshold, i.e., should be optimized as long as this threshold is not achieved.
- A hard bound is a constraint of the form $a \leq dp_i \leq b$, where a, b are constants and dp_i is the i^{th} design parameter.

C-O uses the min/max optimization criterion:

$$\min_{x \in \Omega} (\max_{1 \leq i \leq m} \alpha_i f_i(x)), 0 \leq \alpha_i, \alpha_i \text{ real} \quad (1.5)$$

where $f_i(x)$ is the i th specification and the α_i are user-specified weighting coefficients. The advantage of this formulation is that the optimal value of x can be

placed anywhere in the region of the parameter space bounded by the minima of the individual criteria by appropriate choice of the α_i .

The optimization process is divided into three different phases.

Phase 1:

Not all hard constraints are satisfied. Equation 1.4 takes the form

$$\min_x \max_k \text{hard}_k^s(x) \quad (1.6)$$

$$\text{subject to : } \text{hard_bound}_l(x) \leq 0 \quad \forall l \quad (1.7)$$

Phase 2:

All hard constraints are satisfied (have values ≤ 1). Not all objectives and soft constraints are Level 1. Equation 1.4 takes the form

$$\min_x \max_{i,j} \{\text{obj}_i^s(x), \text{soft}_j^s(x)\} \quad (1.8)$$

$$\text{subject to : } \text{hard}_k^s(x) \leq 0 \quad \forall k \quad (1.9)$$

$$\text{hard_bound}_l(x) \leq 0 \quad \forall l \quad (1.10)$$

Phase 3:

All hard constraints and soft constraints are satisfied and all objectives are Level 1. Equation 1.4 takes the form

$$\min_x \max_i \text{obj}_i^s(x) \quad (1.11)$$

$$\text{subject to : } \text{soft}_j^s(x) \leq 0 \quad \forall j \quad (1.12)$$

$$\text{hard}_k^s(x) \leq 0 \quad \forall k \quad (1.13)$$

$$\text{hard_bound}_l(x) \leq 0 \quad \forall l \quad (1.14)$$

The active objective (constraint), i.e. the one C-O is working on, is displayed in the title of the Pcomb chart in C-O and has a name according to the optimization phase. In Phase 1 it is named MAX_HARD, in Phase 2 it is named MAX_COST_SOFT, and in phase 3 it is named MAX_COST. The value of the displayed active objective (constraint) is scaled so that 1 shows the Level 1/2 border and 2 shows the Level 2/3 border. MAX_HARD and MAX_COST_SOFT are always positive, because the active constraint is Level 3 or Level 2, while MAX_COST is always negative because the active objective is Level 1.

Part III

Specific Work

Chapter 2

F-14 Problem

2.1 Stability and Control Augmentation Systems

Most modern commercial and military aircraft have Stability and Control Augmentation Systems (SCAS) in order to meet the flying qualities requirements. SCAS decrease the pilot workload by controlling the position and angular rate of the aircraft to match the values of pilot input references. The system also reduces the coupling that normally exists between different pilot inputs. The SCAS outputs add to the pilot's. SCAS outputs move only the control surfaces and not the cockpit controls. Automatic pilots, which replace the human pilot when they are in use, are expected to move the cockpit controls.

In the case of high-performance military aircraft, where the pilot may have to maneuver the aircraft to its performance limits and simultaneously perform tasks such as precision tracking of targets, specialized control augmentation systems are needed in addition to the stability augmentation systems [13].

A pitch-axis control augmentation system is a specialized SCAS and it is

used in several modes of operation corresponding to the aircraft type and the required performance.

For a fighter aircraft, where high maneuverability or “agility” counts the most, the suitable controlled variable for the pitch axis is normal acceleration (or load factor) on the aircraft. This is the component of acceleration measured by an accelerometer in the negative direction of the aircraft z-axis [14]. It is directly relevant to performing a maximum-rate turn and must be controllable up to the structural limits of the airframe, or the pilot’s physical limits.

A second common mode of operation for a pitch-axis control augmentation system is as a pitch-rate command system. Control of pitch rate is used when a situation requires precision tracking of a target and is also preferred for approach and landing.

These two modes are contradictory. A control system that has a good normal acceleration step response may have a pitch-rate response with a very large overshoot, and conversely, a reduction in the pitch-rate overshoot may lead to a sluggish normal acceleration response. For fighter aircraft air-combat modes a pitch-rate overshoot is required for good gross acquisition of targets and a dead-beat pitch-rate response is required for good fine tracking. Sometimes the two control schemes are blended together to give the pilot control over pitch rate at low speed and normal acceleration at high speed.

There exists also a third mode of operation, and this is as an attitude (or angle-of-attack) command system. This is the one used for the Grumman F-14 aircraft and will be presented in the following sections.

2.2 Description

In one functional mode, the F-14 uses an angle-of-attack command system for the pitch-axis control augmentation system.

The F-14 problem is to design a system which controls the pitch response to pilot stick commands. The controller performance is tested for speed, angle-of-attack tracking when following an α -command, and for disturbance rejection. The stability, the magnitude of the tail rate, the derivative of the normal acceleration, and the shape of the pitch rate are also controlled.

The problem is derived from a master's thesis example [5]. Some transformations changed the block diagram into an equivalent one, Figure 2.1, that matches the Grumman F-14 Benchmark Control Problem formulations [8], [9], [10]. These transformations are:

The pitch-rate lead filter was changed from $\frac{s+dp_2}{s-dp_1}$ to $\frac{s+dp_2}{s+dp_1}$.

The stick prefilter was changed from $\frac{1}{\frac{1}{dp_7}s+1}$ to $\frac{1}{dp_7s+1}$.

The α -sensor low-pass filter was filter changed from $\frac{1}{\frac{1}{dp_3}s+1}$ to $\frac{1}{dp_3s+1}$.

where

dp_1 = negative of the pitch-rate lead filter pole location.

dp_2 = negative of the pitch-rate lead filter zero location.

dp_3 = negative of the inverse of the α -sensor low-pass filter pole location.

dp_7 = negative of the inverse of the stick prefilter pole location.

The controller gains were changed from

$$u = (reference * dp_6 - feedback) \quad (2.1)$$

$$u = (reference - \frac{feedback}{dp_6})dp_6. \quad (2.2)$$

This implies that dp_6 , appears in the forward path multiplying the output of the stick prefilter but has no effect on the pitch-rate and α -feedback gains where

dp_4 = pitch-rate feedback gain

dp_5 = α -feedback gain

dp_6 = gain of the filtered command signal

dp_8 = common integral gain

White noise filtered through the Dryden wind gust model was added. For optimization the EigLcG1, RMS1spc and BnwPiL2 specs were added. The specifications are shown in Figure 2.6.

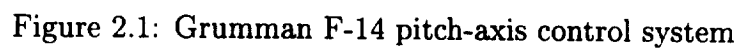
The system has to have the form of Figure 2.1, where parameters dp_1 through dp_8 are to be selected to meet the performance objectives and constraints listed in section 2.4. Selecting the gains dp_1 through dp_8 to yield good closed-loop response to a step input at input 1 and white noise at input 2 corresponds to a multi-input multi-output design problem.

2.3 Block Diagram

Figure 2.1, Figure 2.2, and Figure 2.3 show the block diagram of a pitch-axis control augmentation system. The aircraft dynamics represent those of the Grumman Aerospace F-14 flying at 35,000 ft in level flight at a velocity of 690 ft/sec. The model is linearized, time invariant and continuous time.

The components of the model are:

- short-period longitudinal aircraft dynamics
- tail surface actuator



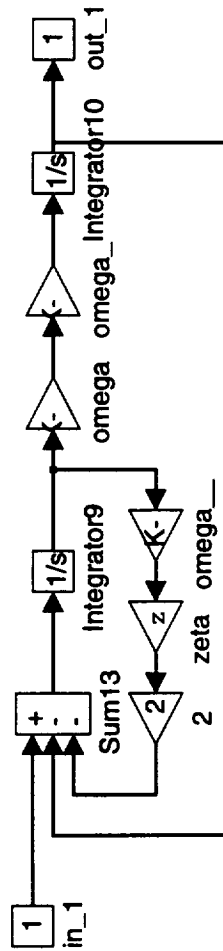


Figure 2.2: Alpha response model

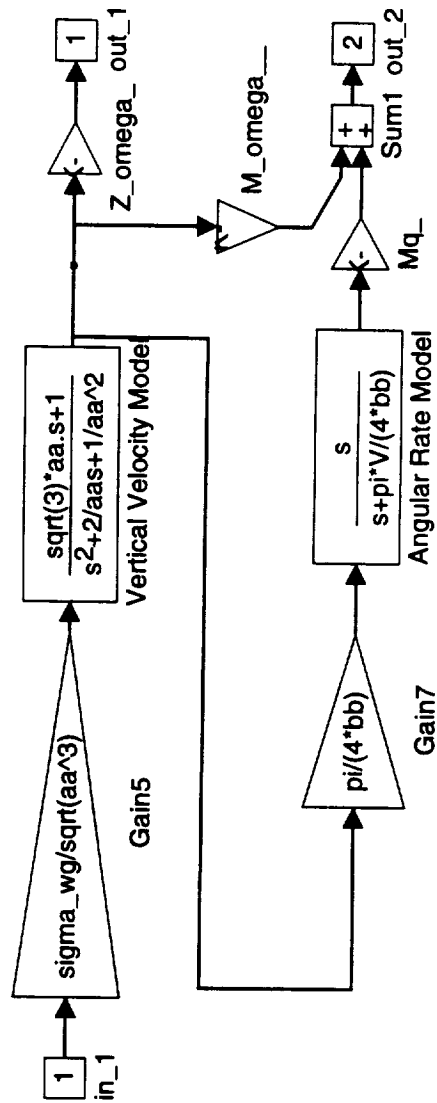


Figure 2.3: Dryden wind gust model

- control system, consisting of a stick prefilter, an error integrator, pitch-rate lead filter, and an angle-of-attack filter
- Dryden wind-gust model

An error integrator has been included to make the control system Type 1, thus ensuring that the aircraft will hold a zero angle-of-attack trajectory when no pressure is applied to the control stick. The Dryden Model filters white noise to form the vertical gust velocity, ω_{gust} , which is then filtered further to give the angular rate gust component, q_{gust} .

The fixed parameter values shown in Figure 2.1 are:

Tail Servo

$$\tau_a = 0.05\text{sec}$$

Aircraft Dynamics

$$Z_\delta = -63.9979 \frac{\text{ft}}{\text{rad} \cdot \text{sec}^2}$$

$$M_\delta = -6.8847 \frac{1}{\text{rad} \cdot \text{sec}^2}$$

$$U_0 = 689.4 \frac{\text{ft}}{\text{sec}}$$

$$M_\omega = -0.00592 \frac{1}{\text{ft} \cdot \text{sec}}$$

$$Z_\omega = -0.6385 \frac{1}{\text{sec}}$$

$$M_q = -0.6571 \frac{1}{\text{sec}}$$

α Response Model

$$\zeta = 0.707$$

$$\omega_0 = 2.49 \frac{\text{rad}}{\text{sec}}$$

Grumman also supplied values for the design parameters dp_1 through dp_8 . These were the results of their quadratic-based optimization system, CASCADE.

These values of:

$$\begin{aligned} dp_1 &= 4.14400e + 00 \frac{\text{rad}}{\text{sec}} & dp_5 &= 6.77000e - 01 \\ dp_2 &= 2.97100e + 00 \frac{\text{rad}}{\text{sec}} & dp_6 &= -1.74600e + 00 \\ dp_3 &= 3.95900e - 01 \text{sec} & dp_7 &= 1.00000e - 01 \text{sec} \\ dp_4 &= 8.15600e - 01 & dp_8 &= -3.86400e + 00 \end{aligned}$$

were used as the initial values for the independent design parameters dp_1 through dp_8 .

2.4 Performance Objectives and Constraints

The second major component, after the system to be controlled and the controller structure, of a control system design problem is the set of specifications. CONDUIT provides a very large collection of aircraft control system specifications in its on-line specs catalog. All the specifications are illustrated in Fig. 2.6.

Specifications are described in an abbreviated fashion. The purpose of the spec is given general terms. This is followed by the CONDUIT description of the spec, written in bold text and followed by the word, “code.” The following lines are written in pseudo-code, with comments. The first line after the header indicates whether the spec applies to the open-loop or closed-loop system. This same line specifies the inputs and outputs used in computing the spec. The next line, in the absence of comments, designates the spec as an objective, hard constraint or soft constraint. This line also indicates the range of frequencies or times over which the spec is evaluated. The last line gives the goal the spec is intended to achieve. Pictures of all the specs can be found in Fig. 2.6.

Performance Objective

This specification, when used as a performance objective, generally tries to make the closed-loop bandwidth larger and the time delay smaller.

1. The pitch bandwidth specification requires a lower limit for pitch attitude bandwidth and an upper limit for time delay. Aircraft maneuvers are classified in different flight phase categories, corresponding to the required speed of the maneuvers, precision of the tracking, precision of the flight-path control [20]. This spec applies to all categories so it is specified as Categories A&D.

BnwPiL2 code:

closed-loop, input 1 (the pilot pitch command), output 7 (α (pitch response)).

objective, $\omega \in [0.1, 20]$ Hz

maximize bandwidth and minimize time delay

Constraints

1. The first stability criterion of the total feedback system (i.e., with both feedback paths active) requires nonpositive eigenvalues:

EigLcG1 code:

closed-loop, from any input to any output

[a,b,c,d]=closed loop model, *unstab* equals sum of positive eig(a) if there are some positive eigenvalues or *unstab* equals max of negative eig(a) if all eigenvalues are negative

hard constraint

minimize *unstab* : Level 1/2 = 0.001, Level 2/3 = 0.002

2. The second stability criterion of the system is:

gain margin ≥ 6 [dB]

phase margin ≥ 45 [deg]

StbMgG1 code :

open-loop, input 3 broken-loop input, output 6 broken-loop output

hard constraint, $\omega \in [0.1, 100]$ Hz

maximize gain : Level 1/2 = 6 [dB], Level 2/3 = 4 [dB]

maximize phase: Level 1/2 = 45 [deg], Level 2/3 = 35 [deg]

3. The angle of attack, α [deg], response to a pilot-step input of 2.0 [deg] is required to match that of a given critically damped second order system. The resulting difference is to be minimized with an acceptable value being 0.2 [deg]. Because the system is linear this is equivalent to acceptable design for α error less than 10%, satisfactory design for α error less than 5%.

aerrspc code:

closed-loop, input 1 the pilot pitch command, output 2, q =pitch rate

soft constraint, $t \in [0.1, 3]$ sec

minimize α error, Level 1/2 = [-0.05 0.05], Level 2/3 = [-0.1 0.1]

4. The maximum velocity of the tail surface deflection servo must be less than 25.0 [deg/sec] for the 2.0 [deg] pilot-step input. Because the system is linear, this is equivalent to satisfactory design for ratio between tail rate and command less than 12.5%.

tailspc code:

closed-loop, input 1 the pilot pitch command, output 1 the tail rate

soft constraint, $t \in [0.05, 3]$ sec

minimize tail rate, Level 1/2 = [-12.5 12.5], Level 2/3 = [-15 15]

5. The rate of change of acceleration (i.e. jerk) experienced by the pilot in the vertical direction must be greater than zero during the response to the pilot-step input. The normal acceleration is defined as

normal acceleration = $U_0 * q + L * \dot{q} - \dot{\omega}$, where U_0 appears in Figure 2.1, at aircraft dynamics, q is the pitch rate, \dot{q} is the derivative of the pitch rate, $\dot{\omega}$ is the output from Sum6 in Figure 2.1, and $L = 22.8$ [ft]. The derivative of normal acceleration is obtained from the normal acceleration by a multiplication by $\frac{10s}{s+10}$ in the block diagram. The derivative of normal acceleration is scaled by the acceleration due to gravity, $g_0 = 32.2 \frac{ft}{sec^2}$.

For this spec two signals were considered, Signal 1 corresponding to an input of 2 [deg] and Signal 2 corresponding to an input of 5.72 [deg].

This constraint must be positive, but the value does not matter. Thus a saturation was introduced in the block diagram, with limits ± 1 . If the value is smaller than -1, the spec will show -1, which is Level 3. If the value is bigger than 1, the spec will show 1 and is Level 1. If the saturation were omitted then the envelope spec would need an expanded range for Level 1, and this would decrease the readability of the spec.

jerkspc code:

closed-loop, input 1 pilot pitch command, output 4 the jerk ($\frac{da}{dt}$, a = acceleration) experienced by the pilot

soft constraint, $t \in [0.1, 1.5]$ sec

maximize jerk, Level 1/2 = [0 1] [1/sec], Level 2/3 = [-0.2 1.1] [1/sec],

saturation [-1, 1]

6. The pitch rate (q) [rad/sec] response to stick inputs [rad] is required to fit the lower-order system [6]:

$$\frac{K(1.7s + 1)e^{-\tau s}}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \quad (2.3)$$

within the bounds of

$$2.122 \leq \omega_0 \leq 3.5214$$

$$0.5 \leq \zeta \leq 0.707$$

$$\tau \leq 0.1$$

$$\text{for } 0.2 \leq \omega \leq 20$$

Because K is unconstrained, this requirement constrains only the transient response. Equation 2.3 and the associated bounds can be translated into an equivalent gain and phase requirement.

The **gain requirement** is:

$$\min_{\omega_0, \zeta} \frac{\sqrt{1 + (1.7\omega)^2}}{\sqrt{(1 - (\frac{\omega}{\omega_0})^2)^2 + (2\zeta\frac{\omega}{\omega_0})^2}} \leq \frac{gain}{dcgain} \leq \max_{\omega_0, \zeta} \frac{\sqrt{1 + (1.7\omega)^2}}{\sqrt{(1 - (\frac{\omega}{\omega_0})^2)^2 + (2\zeta\frac{\omega}{\omega_0})^2}} \quad (2.4)$$

$$\text{for } 0.2 \leq \omega \leq 20$$

The minimum is attained for $\omega_0 = 2.122$ and $\zeta = 0.707$. The maximum is attained for $\omega_0 = 3.5214$ and $\zeta = 0.5$. Equation 2.4 is equivalent to

$$\frac{\sqrt{1 + (1.7\omega)^2}}{\sqrt{(1 - (\frac{\omega}{2.122})^2)^2 + (2 * 0.707 \frac{\omega}{2.122})^2}} \leq \frac{gain}{dcgain} \leq \frac{\sqrt{1 + (1.7\omega)^2}}{\sqrt{(1 - (\frac{\omega}{3.5214})^2)^2 + (2 * 0.5 \frac{\omega}{3.5214})^2}} \quad (2.5)$$

$$\text{for } 0.2 \leq \omega \leq 20$$

The **phase requirement (time delay)** is:

$$\min_{\omega_0, \zeta, \tau} \left(\arctan 1.7\omega - \arctan \frac{2\zeta \frac{\omega}{\omega_0}}{1 - \left(\frac{\omega}{\omega_0}\right)^2} - \tau\omega \right) \frac{360}{2\pi} [\text{deg}] \leq \Phi \leq \max_{\omega_0, \zeta, \tau} \left(\arctan 1.7\omega - \arctan \frac{2\zeta \frac{\omega}{\omega_0}}{1 - \left(\frac{\omega}{\omega_0}\right)^2} - \tau\omega \right) \frac{360}{2\pi} [\text{deg}] \quad (2.6)$$

for $0.2 \leq \omega \leq 20$

The minimum is attained for $\omega_0 = 2.122$, $\zeta = 0.707$, and $\tau = 0.1$. The maximum is attained for $\omega_0 = 3.5214$, $\zeta = 0.5$, and $\tau = 0$. Equation 2.6 is equivalent to

$$\left(\arctan 1.7\omega - \arctan \frac{2 * 0.707 \frac{\omega}{2.122}}{1 - \left(\frac{\omega}{2.122}\right)^2} - 0.1 * \omega \right) \frac{360}{2\pi} [\text{deg}] \leq \Phi \leq \left(\arctan 1.7\omega - \arctan \frac{2 * 0.5 \frac{\omega}{3.5214}}{1 - \left(\frac{\omega}{3.5214}\right)^2} \right) \frac{360}{2\pi} [\text{deg}] \quad (2.7)$$

for $0.2 \leq \omega \leq 20$

Another form of this requirement is the “YF-17” frequency-response criterion [16]. It shows the matching of the high order system (HOS) pitch rate to stick input with some gain and phase curves. This criterion includes a constraint for the dcgain. Figures 2.4 and 2.5 show a comparison between (dcgain*LOS) and “YF-17” criterion. For our scheme dcgain=0.5697 at the nominal point. Therefore gain and phase curves of

$$0.5697 * \frac{(1.7s + 1)e^{-\tau s}}{\frac{s^2}{\omega_0^2} + \frac{2\zeta s}{\omega_0} + 1} \text{ and “YF-17”. will be compared.}$$

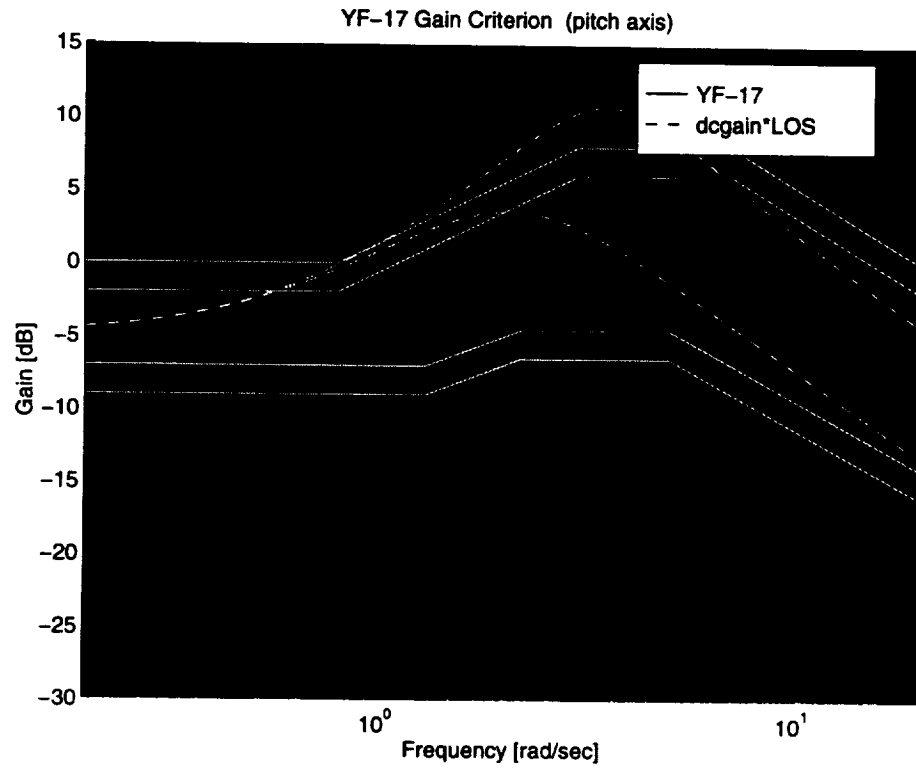


Figure 2.4: Comparison between YF-17 criterion and analytic curve of $\text{dcgain} \cdot \text{LOS}$. This is the gain spec. The YF-17 Level 1 region is indicated in black.

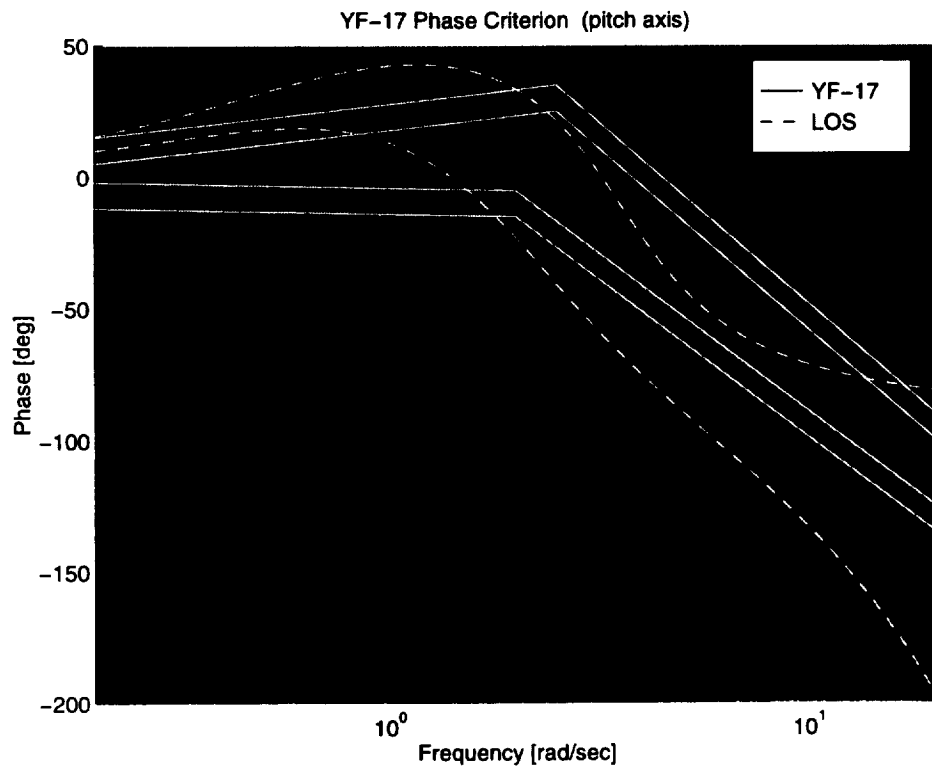


Figure 2.5: Comparison between YF-17 criterion and analytic curve of dc-gain*LOS. This is the phase spec. The YF-17 Level 1 region is indicated in black.

The theoretical criterion, which considers the gain and phase limits for the pitch rate (divided by $dcgain$), will be used in the optimization. The practical criterion “YF-17” will be shown as “check only” in the performance chart, without any role in optimization. The $dcgain$ option for gain is set to $20 \cdot \log_{10}(0.5697) = -4.8871$ [dB].

YFGAspc and YFPHspc code:

closed-loop, input 1 pilot pitch command, output 3 pitch rate

soft constraint, $\omega \in [0.2, 20]$ Hz

7. RMS g's at the pilot caused by turbulence have to be minimized. Disturbance rejection is an important aspect of any control system, so there is a soft constraint which takes care of errors due to random disturbances. Berman and Gran [13] suggest that the sample rate selection for aircraft autopilots should be done based on its effect on disturbance rejection and discuss an application of digital design and sample rate determination to pitch-plane control of the Grumman V/STOL Design 607A.

Disturbances enter a system with various characteristics ranging from steps to white noise. For purposes of determination of sample rate, the higher frequency random disturbances are the most influential; therefore, we will concentrate on their effect. In other words, disturbances that are fast compared to the plant and the sample rate will be compared. That is, where the noise can be considered to be white.

In case the control system uses a good continuous controller, the magnitude of the error response represents a lower bound on the magnitude of the error response that can be hoped for when implementing the controller

digitally. Some degradation from the continuous design occurs because of the sampling. In order to analyze the degradation of the digital controller as compared to the continuous controller, it is important to consider the effect of the noise consistently, with both the continuous and the digital controllers.

Suppose the system is continuous and represented by

$$\dot{x} = Ax + Bw \quad (2.8)$$

where system matrices A and B describe a closed-loop system including a continuous controller and $w = [w_1 w_2 \dots w_n]^T$ is the noise. Denote the power spectral density of w as R_{wpsd} (alternatively referred to as the “white-noise intensity” or “mean-square spectral density”). Then the covariance matrix of w is $E[w(t)w^T(t + \tau)] = R_{wpsd}\delta(\tau)$.

This assumes that the disturbance $w(t)$ is white noise, which means that $w(t)$ and $w(s)$ are uncorrelated for all $t \neq s$.

Typically, if there is more than one process noise component ($i > 1$), one has no information on the cross-correlation of the noise elements and therefore R_{wpsd} is selected as a diagonal matrix.

The steady state value of the covariance matrix of x is given by the Lyapunov equation

$$AX + XA^T + BR_{wpsd}B^T = 0 \quad (2.9)$$

The solution to this equation, $X = E[x(t)x^T(t)]$ represents the amplitude of the random response of the state due to the excitation from w . X can be used to establish a baseline against which controllers are com-

pared. For this particular problem the root mean square (RMS) value, computed as the square root of the steady state value of the covariance of y , $\sqrt{E[Cx(t)x^T(t)C^T]} = \sqrt{CXC^T}$ will be used. y is a scalar because one particular output is considered, e.g. pilot g's. And R_{wpsd} is assumed to be 1.

The maximum RMS(root mean squared) must be less than 0.04 [8,9].

RMS1spc code:

closed-loop, input 2 white noise input to the Dryden wind gust model,
output 5 pitch acceleration (g's)

soft constraint

minimize RMS, Level 1/2 = 0.04, Level 2/3 = 0.1 where

$$X = \text{lyap}(A, B(:, 2) * B(:, 2)^T)$$

$$gVar = C(5, :) * X * C(5, :)^T$$

$$RMS = \sqrt{gVar}$$

It is also necessary to evaluate X when the system has a digital controller for the identical excitation applied to the plant. Suppose the discrete equivalent system is

$$x(k+1) = A_d x(k) + B_d w(k) \quad (2.10)$$

where the process noise $w(k)$ is a random sequence with zero mean, that is, $E[w(k)] = 0$, and has no time correlation or is "white noise",

$E[w(i)w^T(j)] = 0$ if $i \neq j$, and has covariances or "noise levels" defined by

$E[w(k)w^T(k)] = R_w$. The desired result, called the discrete Lyapunov equation, is

$$A_d X A_d^T + G = X \quad (2.11)$$

where $G = \int_0^T A_d(\tau) B R_{\text{wpsd}} B^T A_d^T(\tau) d\tau$. If T , the sampling interval, is much shorter than all system time constants, that is, $A_d \cong I$ and $B_d \cong BT$, then the integral is approximately $G \cong B_d \frac{R_{\text{wpsd}}}{T} B_d^T = B_d R_w B_d^T$ [12]. In order to evaluate the effect of sample rate on the performance of a controller in the presence of white plant disturbances, first Eq.2.9 should be evaluated to find the baseline covariance X and then Eq.2.11 should be repeatedly evaluated with varying sample rates to establish the degradation versus sampling. The RMS value is the quantity that is typically measured. In the discrete case, for the F-14 particular problem, the RMS value is computed in MATLAB as follows:

$$X = \text{dlyap}(A_d, B_d(:, 2) * B_d(:, 2)^T / T)$$

$$\text{gVar} = C_d(5, :) * X * C_d(5, :)^T$$

$$\text{RMS} = \sqrt{\text{gVar}}$$

Taking relatively fast sampling (twenty times the fastest frequency mode), the RMS value is almost the same for both the continuous and digitized system, so that either one can be used, in RMS computation. In case that $dt=1/50$ and nominal design parameters are used $\text{rms}_{\text{continuous}} = 0.0367244$ and

$$\text{rms}_{\text{discrete}} = 0.0367230.$$

2.5 Setup of the Initialization File

Looking at the closed-loop eigenvalues at the nominal starting point, the fastest pair of eigenvalues is $-9.84 \pm 9.57j$, giving the largest $\omega = 13.72 \frac{\text{rad}}{\text{sec}}$, and the largest $\nu = 2.18 \text{ Hz}$. The sampling rate is chosen to be 50 samples per second.

epsilonF was set at 0.0002 (the default value) and 0 (machine precision), and tests were made for both values.

The input signals InpSig1 and InpSig2 are step signals, corresponding to 2 and 5.72 [deg] respectively.

2.6 Optimization

The F-14 optimization starting at two different initial points will be presented: a satisfactory starting point found by Grumman using their quadratic-based optimization system CASCADE, called the nominal point, and an arbitrarily chosen point, called the trivial point given below:

$$\begin{aligned} dp_1 &= 1 & dp_5 &= 1 \\ dp_2 &= 1 & dp_6 &= -1 \\ dp_3 &= 1 & dp_7 &= 1 \\ dp_4 &= 1 & dp_8 &= -1 \end{aligned}$$

In both cases CONDUIT was able to improve the performance and to deliver a satisfactory result. The optimization results using two values for epsilonF are shown in Table 2.1.

The optimal point achieved when the optimization started in the nominal point and used $\epsilon_F = 0$ was:

$$\begin{aligned} dp_1 &= 4.14184e + 00 & dp_5 &= 6.77324e - 01 \\ dp_2 &= 2.97359e + 00 & dp_6 &= -1.74871e + 00 \\ dp_3 &= 3.93938e - 01 & dp_7 &= 5.28613e - 02 \\ dp_4 &= 8.22272e - 01 & dp_8 &= -3.86422e + 00 \end{aligned}$$

Table 2.1: Comparison between the optimal points obtained by CONDUIT for the F-14 problem

	Starting point	Optimal using epsilonF=0	Optimal using epsilonF=0.0002
Nominal	<i>MAX_COST</i> -0.503495	<i>MAX_COST</i> -1.2329	<i>MAX_COST</i> -0.955191
Trivial	<i>MAX_COST_SOFT</i> 19.6509	<i>MAX_COST_SOFT</i> 0.0779831	<i>MAX_COST_SOFT</i> 0.376201

The optimal point achieved when the optimization started in the nominal point and used $\text{epsilonF} = 0.0002$ was:

$$\begin{aligned}
 dp_1 &= 4.19590e + 00 & dp_5 &= 6.84997e - 01 \\
 dp_2 &= 3.03156e + 00 & dp_6 &= -1.69662e + 00 \\
 dp_3 &= 4.25392e - 01 & dp_7 &= 8.74300e - 02 \\
 dp_4 &= 8.41695e - 01 & dp_8 &= -3.87025e + 00
 \end{aligned}$$

The largest percentage variation between these two optimal points is 65% for dp_7 , and the largest variation is 0.0580 for dp_2 . These solutions are close one to another.

The optimal point achieved when the optimization started in the trivial point and used $\text{epsilonF} = 0$ was:

$$\begin{aligned}
 dp_1 &= 8.00974e - 01 & dp_5 &= 4.72042e - 01 \\
 dp_2 &= 1.34126e + 00 & dp_6 &= -3.33092e + 00 \\
 dp_3 &= 1.30245e + 00 & dp_7 &= 2.20404e - 01 \\
 dp_4 &= 5.26028e - 01 & dp_8 &= -1.45150e + 00
 \end{aligned}$$

The optimal point achieved when the optimization started in the trivial point and used $\epsilon_F = 0.0002$ was:

$$\begin{aligned} dp_1 &= 9.57795e - 01 & dp_5 &= 7.17266e - 01 \\ dp_2 &= 9.82241e - 01 & dp_6 &= -1.70196e + 00 \\ dp_3 &= 8.40876e - 01 & dp_7 &= 4.88951e - 02 \\ dp_4 &= 5.56358e - 01 & dp_8 &= -1.07892e + 00 \end{aligned}$$

The largest percentage variation between these two optimal points is 350% for dp_7 , and the largest variation is 1.6290 for dp_6 .

It can be seen in Table 2.1 that for the F-14 problem the results were better when $\epsilon_F=0$ was used. The performance is presented in Figures 2.6-2.10,

In Figure 2.7 it can be seen that the objective function, pitch bandwidth and time delay were improved at the expense of the tail rate soft constraint, which was pushed to its border. The speed of the airplane response cannot increase more, because this would violate the limit for the tail rate. Note that the YF-17 spec is check only. Thus, CONDUIT ignores the fact that the phase portion of the spec is at Level 3.

In Figure 2.8 it can be seen that the objective was improved at the expense of the attitude tracking soft constraint (Alpha error), which was pushed to its limit. The previous comment on the YF-17 spec applies here as well.

In Figure ?? it can be seen that at the optimal point all objectives and constraints are Level 1 excepting two of them (Alpha error and Gain Requirement) which exceed insignificantly the Level 1/2 border. Notice particularly that the theoretical model-following spec is satisfied. It is this spec that was used, not the YF-17, which is check only.

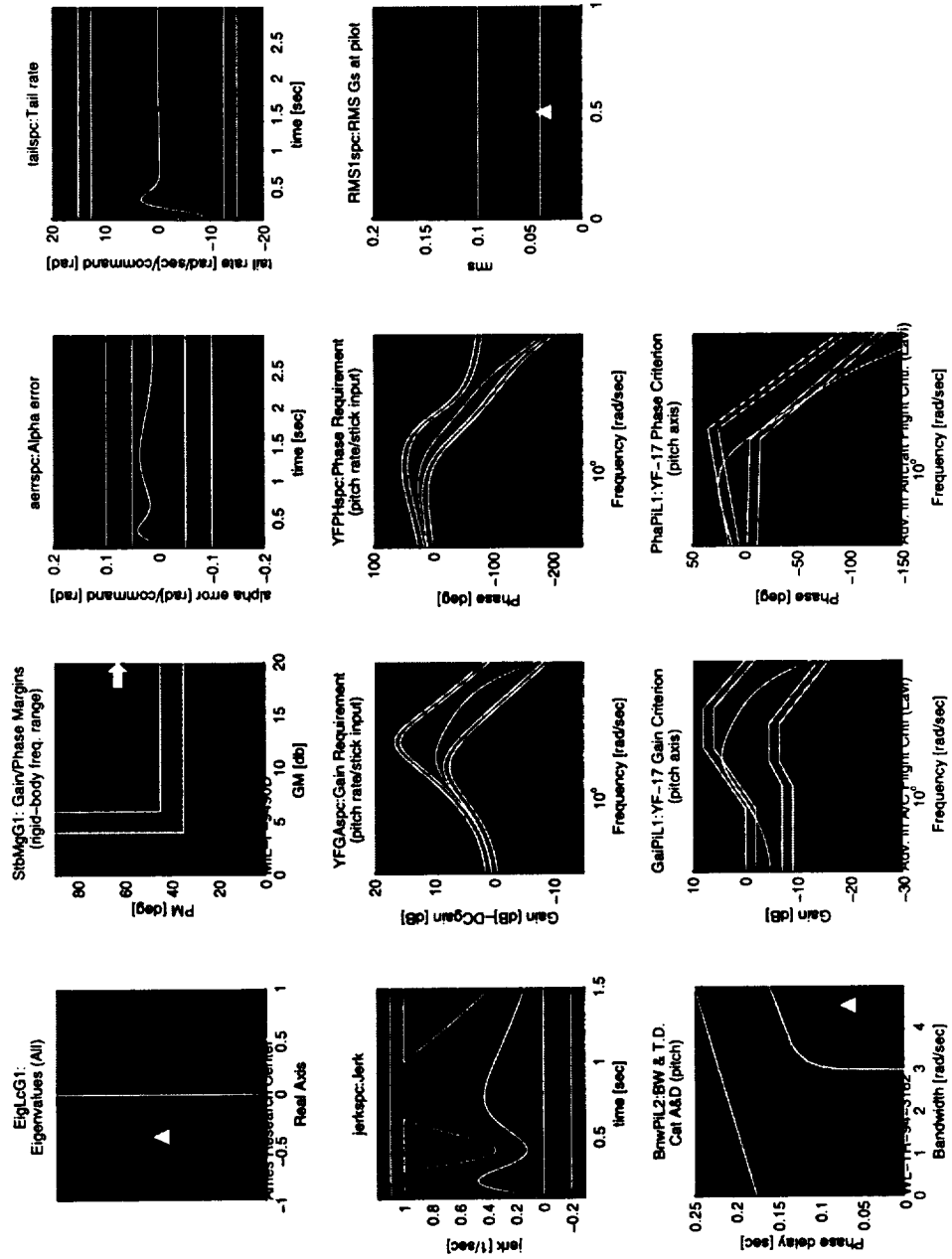


Figure 2.6: Grumman F-14, nominal starting point

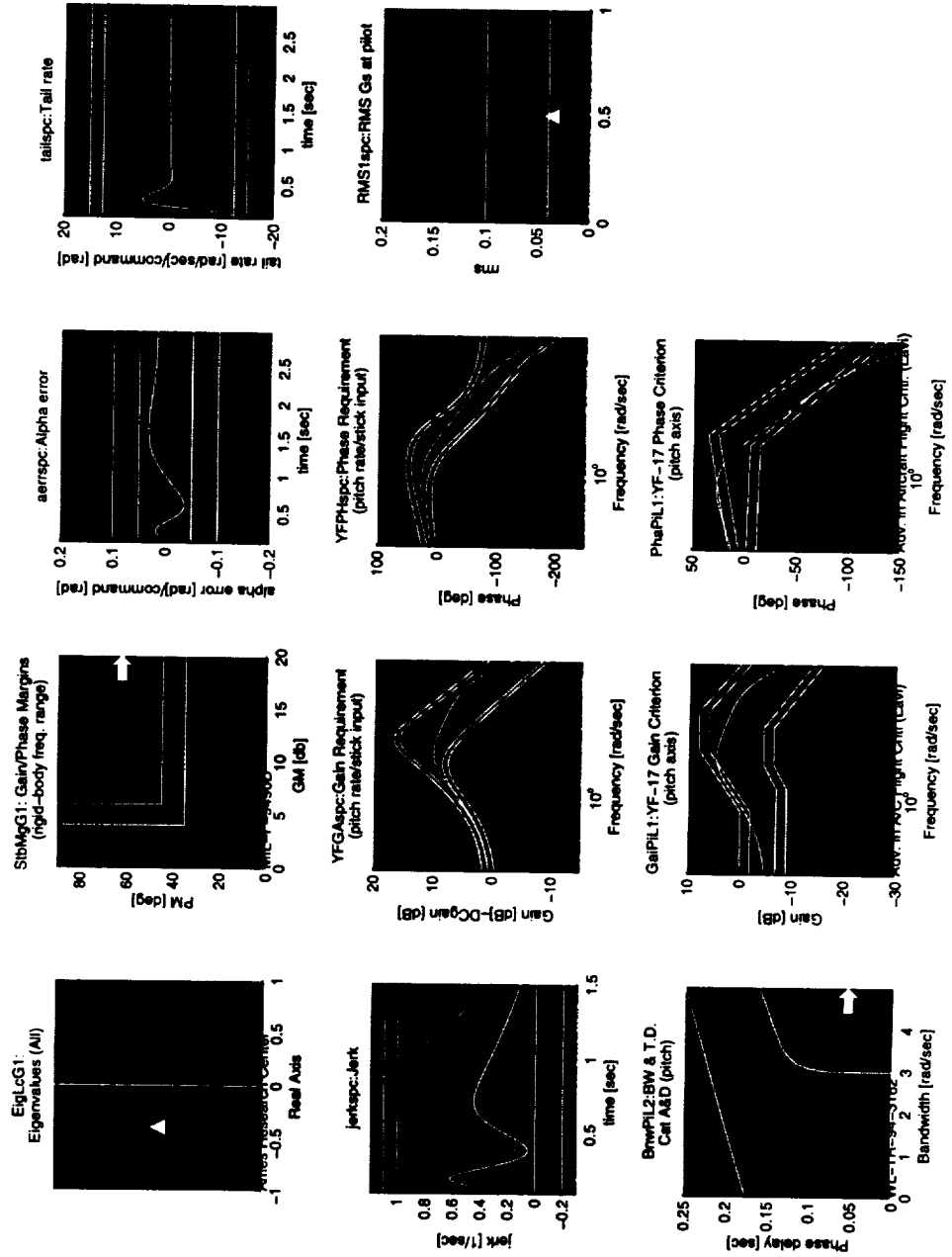


Figure 2.7: Grumman F-14, nominal point improved by CONDUIT, $\epsilon_F=0$

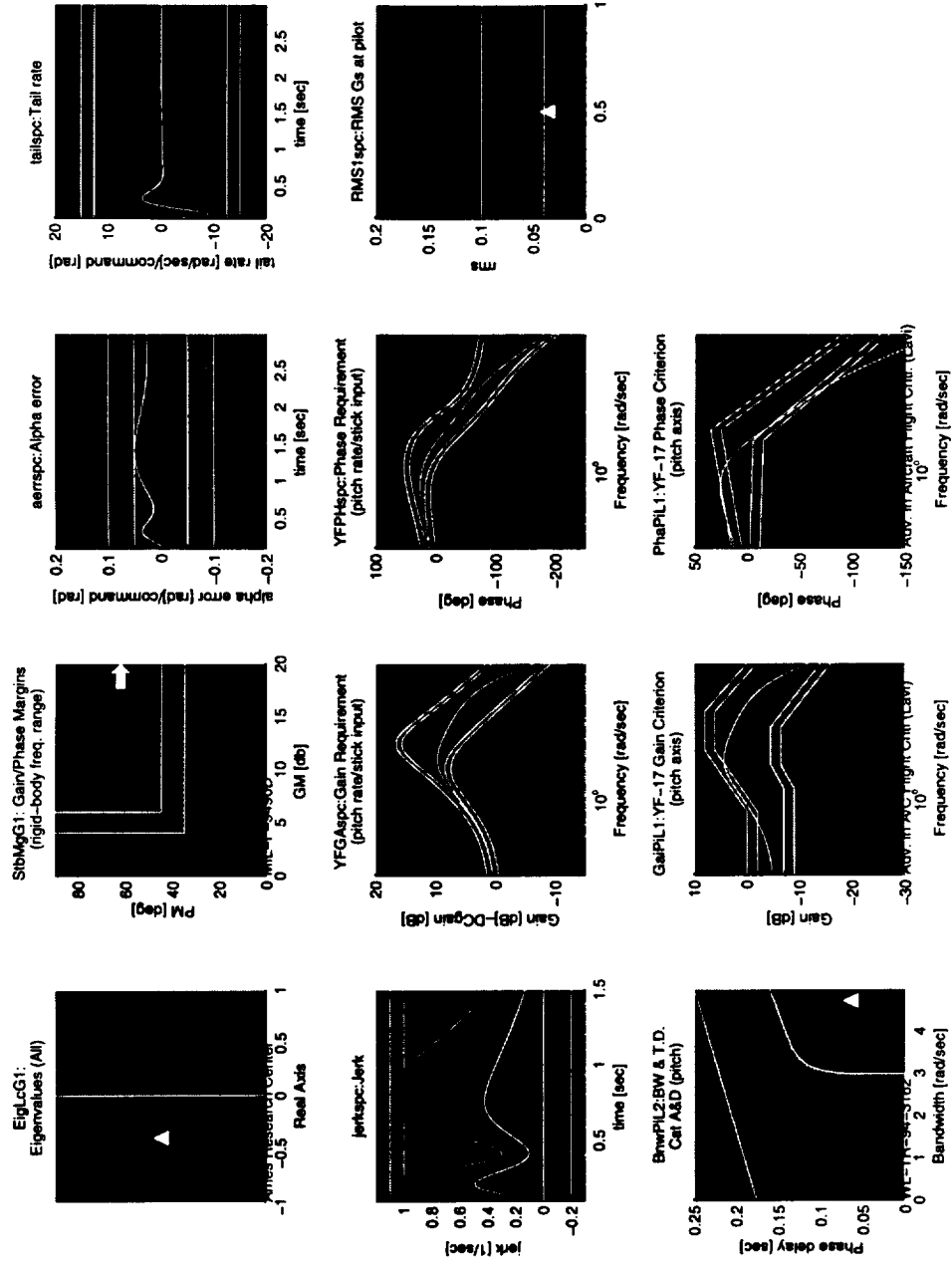


Figure 2.8: Grumman F-14, nominal point improved by CONDUIT, ϵ_{psi}
lonF=0.0002

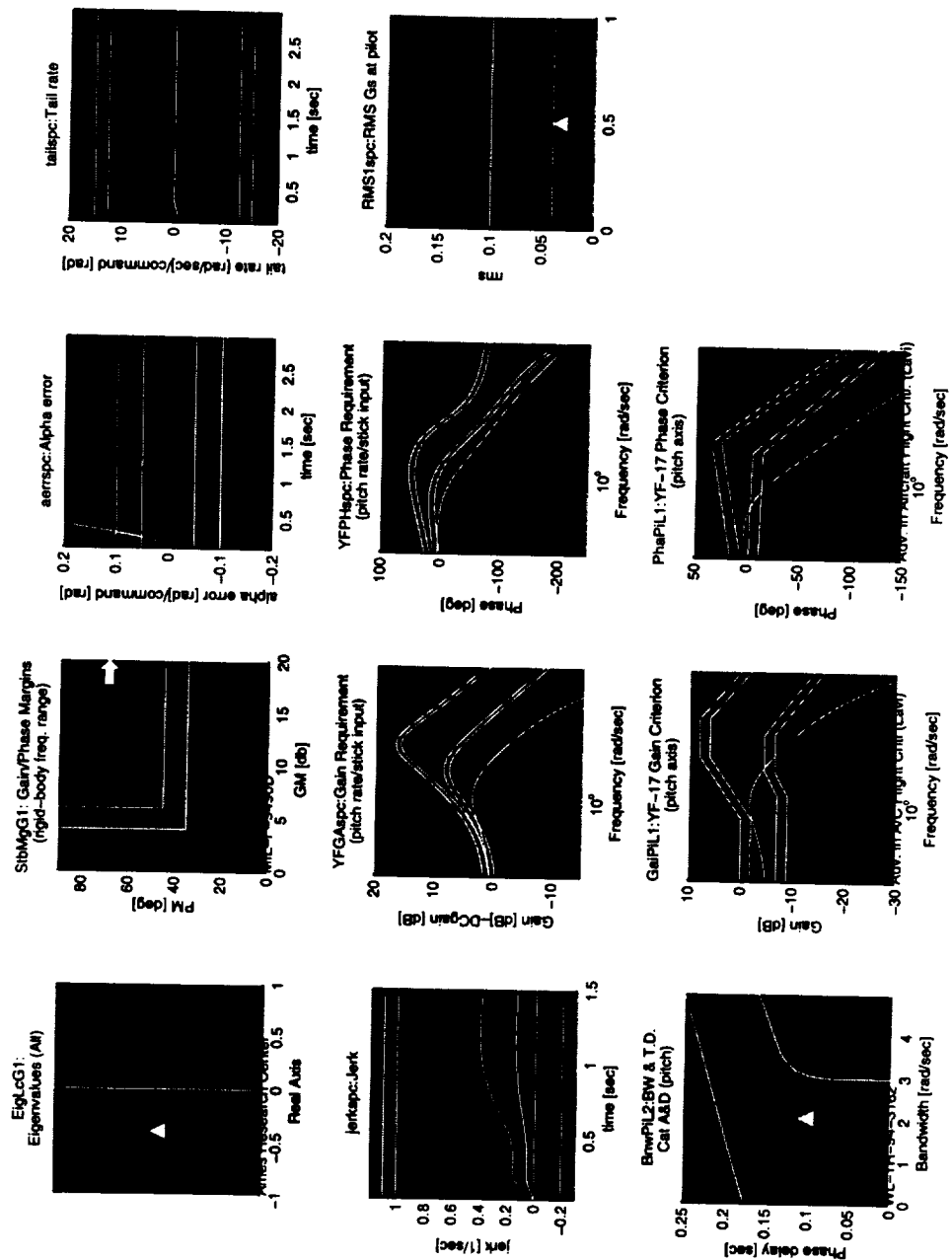


Figure 2.9: Grumman F-14, trivial starting point

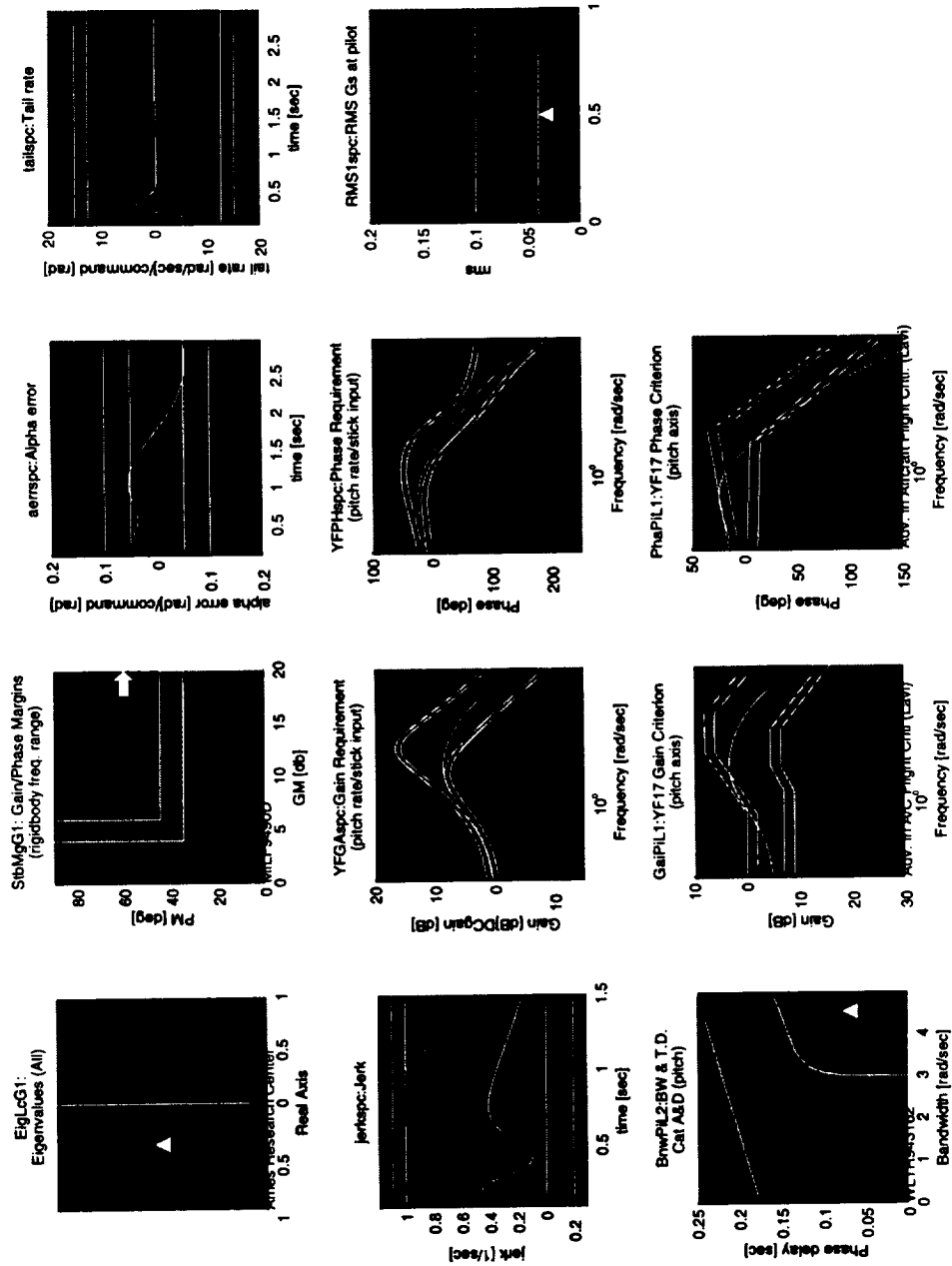


Figure 2.10: Grumman F-14, trivial point improved by CONDUIT, epsilonF=0

Chapter 3

Antenna Problem

3.1 Description

Usually a controller for an antenna is tested for tracking errors when following a command and for servo errors due to disturbances. The tracking and disturbance rejection properties have to be traded off. The feedback loop is effective in suppressing the perturbation, but has weaker tracking properties. On the other hand the feedforward controller has good tracking properties, confirmed by both simulations and field measurements, but its ability to compensate the perturbations is insufficient. When comparing the properties of feedback and feedforward, one can conclude that by combining the two it is possible to improve both tracking and disturbance rejection properties.

This problem is derived from a master's thesis example, [5]. The problem is to design a servomechanism for controlling the line-of-sight (LOS) of an airborne electro-optical device about a particular axis of rotation. The controller's most critical function is to isolate the LOS from perturbations induced by aircraft angular motions. So, only the feedback loop that assures perturbation rejection

will be discussed in the following. Aircraft angular motions produce disturbance torques which must be nullified by the servomechanism's torque motor. The torque motor is driven from feedback elements mounted on the controlled device. These elements, a rate-measuring gyroscope, and an angular accelerometer are inertial in character and hence measure LOS in an absolute sense.

Unlike the F-14 problem, the concern of this design is not the precision with which the system responds to command inputs. Rather, the goal here is to design a system whose response to disturbance inputs is minimized. The reader may find other instructive contrasts to the F-14 problem:

1. As is typical in the servomechanisms field, where design via the time domain methods of modern control theory has never really taken hold, the design criteria are specified completely in the frequency domain.
2. Because higher frequencies are of interest here, more so than in the F-14 problem, the design must consider dynamic characteristics of the actuator, mechanical structure, and feedback transducers themselves. These become part of the system "plant". These issues are also important in aircraft SCAS and would be included in a more realistic SCAS design problem.
3. Because of the above and the corresponding increase in the complexity of the specified form of the controller, the state dimension of the control system is significantly higher.
4. Because the magnitude of the design parameters covers a much larger range, the problem is normalized by use of the "variation" specification, a feature of C-0 [1, p1.6] that is not presently implemented in CONDUIT.

3.2 Block Diagram

Figure 3.1 is the block diagram of an antenna control system.

The components of the model are:

- antenna dynamics
- power amplifier and torque motor
- angular accelerometer and accelerometer filter
- rate gyroscope and gyroscope filter
- control system, consisting of a lead-integrator, quadratic low-pass filter and quadratic lead-lag for the rate loop and lead-integrator and quadratic lag-lead for the acceleration loop

The fixed parameter values of the plant shown in Figure 3.1 are:

$km = 0.801 \frac{\text{oz}\cdot\text{in}}{\text{v}}$ power amplicator and motor gain

$fm = 796 \text{ Hz}$ motor electrical bandwidth

$J = 0.11 \text{ oz}\cdot\text{in}\cdot\text{s}^2$ gimbal inertia

$fr = 160 \text{ Hz}$ gimbal resonance frequency

$far = 122 \text{ Hz}$ gimbal anti-resonance frequency

$Zr = 0.015$ gimbal structural zeta

$asf = 0.291 \frac{\text{v}}{\frac{\text{rad}}{\text{s}^2}}$ accelerometer scale factor

$fabw = 350 \text{ Hz}$ accelerometer bandwidth

$faf = 500 \text{ Hz}$ accelerometer filter bandwidth

$rsf = 2.263 \frac{\text{v}}{\frac{\text{rad}}{\text{s}}}$ gyro scale factor

$frbw = 75 \text{ Hz}$ gyro natural frequency

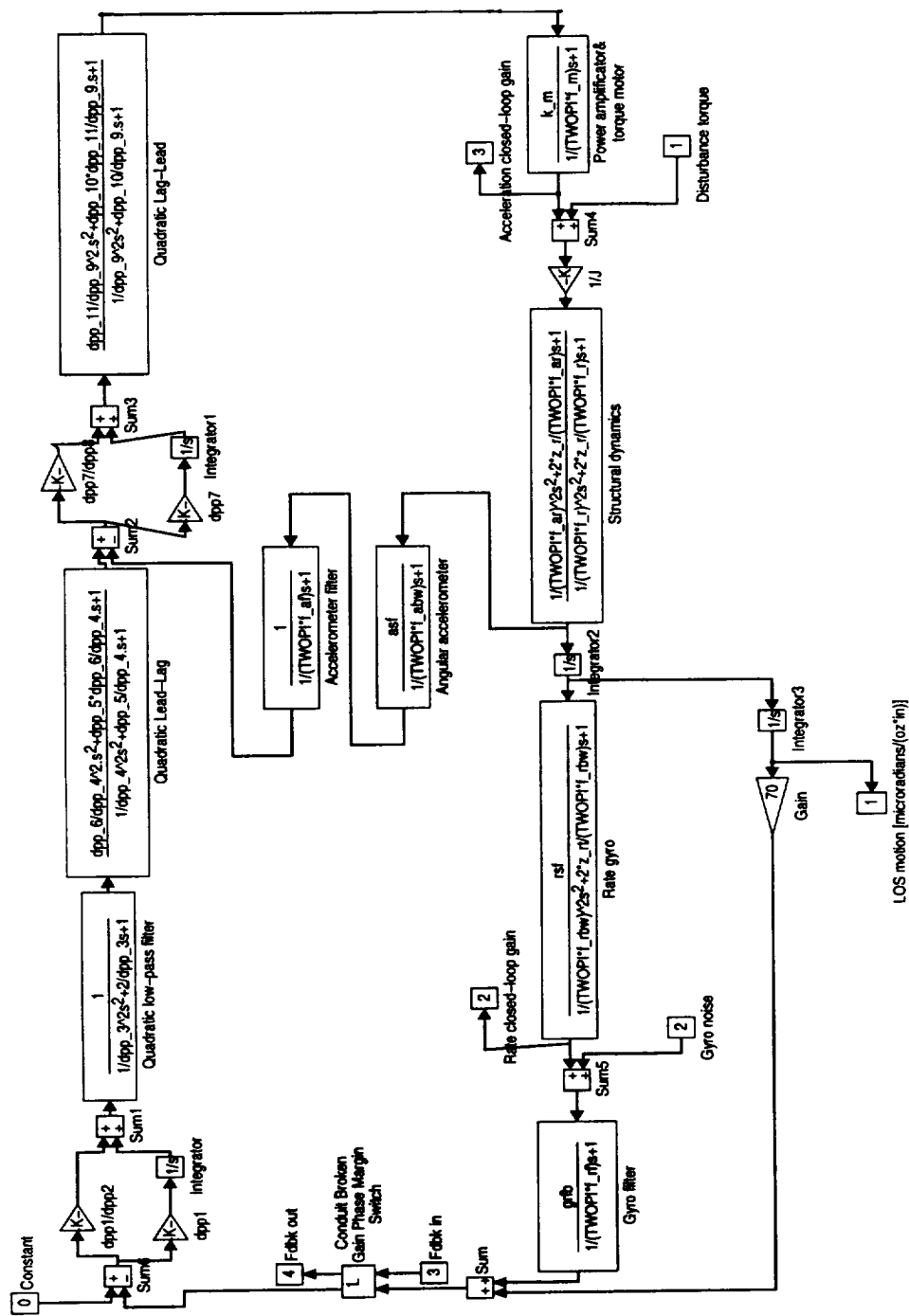


Figure 3.1: Antenna control system

$$zrt = 0.7 \text{ gyro zeta}$$

$$grfb = 10 \frac{v}{v} \text{ gyro filter gain}$$

$$frf = 482 \text{ Hz gyro filter bandwidth}$$

There are eleven parameters that have to be chosen to meet the desired specifications:

$$k1 = \text{rate-loop integral gain}$$

$$f1 = \text{negative of rate-loop lead pole}$$

$$f2 = \text{rate-loop second-order filter bandwidth}$$

$$f3 = \text{rate-loop second-order lead-lag filter lag frequency}$$

$$f4 = \text{negative of the acceleration-loop lead pole}$$

$$f5 = \text{acceleration-loop second-order lag-lead lag frequency}$$

$$z1 = \text{rate-loop second-order lead-lag filter } \zeta$$

$$z2 = \text{acceleration-loop second-order lag } \zeta$$

$$a1 = \text{rate-loop second-order lead-lap filter}$$

$$a2 = \text{acceleration-loop second-order lag-lead filter}$$

In order to simplify the block diagram the design parameters used by CONDUIT are formed from the previously listed parameters as indicated below:

$$dpp_1 = k1$$

$$dpp_2 = 2 * pi * f1$$

$$dpp_3 = 2 * pi * f2$$

$$dpp_4 = 2 * pi * f3$$

$$dpp_5 = 2 * z1$$

$$dpp_6 = a1^2$$

$$dpp_7 = k2$$

$$dpp_8 = 2 * pi * f4$$

$$dpp_9 = 2 * pi * f5$$

$$dpp_{10} = 2 * z2$$

$$dpp_{11} = a2^2$$

For this problem the design parameters are required to be positive. This insures stability and minimum phase for the lead-lag compensator in the rate loop and for the lag-lead compensator in the acceleration loop.

3.3 Performance Objectives and Constraints

The second part of the formulation of the control problem is the selection of the specifications. The specifications are the same as those in [5] except that two specs have been added. The specs have been rewritten in the appropriate form for CONDUIT. All the specifications are summarized below. They are all illustrated in Fig. 3.2.

Objectives

1. The objective is to minimize the motion of the LOS in response to a disturbance torque input. This minimization is to be achieved over a disturbance frequency range from 0.1 Hz to 1000 Hz. The "good" and "bad" values for this quantity are 60 and 120 microradians/oz * in, respectively.

torqspc code:

closed loop, input 1 disturbance torque, output 1 LOS motion

objective, $\omega \in [10^{-1}, 10^3]$ Hz

minimize, Level 1/2 = $[-60 \ 60][\frac{\text{microradians}}{\text{oz*in}}]$, Level 2/3 = $[-80 \ 120][\frac{\text{microradians}}{\text{oz*in}}]$

Constraints

1. Stability criterion of the total feedback system (i.e., with both feedback paths active) is “all eigenvalues negative”.

EigLcG1 code:

closed loop, from any input to any output

[a,b,c,d]=closed-loop model, *unstab* equals sum of positive eig(a) if there are some positive eigenvalues or *unstab* equals max of negative eig(a) if all eigenvalues are negative

hard constraint

minimize *unstab* : Level 1/2 = 0.001, Level 2/3 = 0.002

2. Stability margins of the feedback system are:

gain margin ≥ 6 [db]

phase margin ≥ 45 [deg]

StbMgG1 code:

open loop, input 3 broken-loop input, output 4 broken-loop output

hard constraint, $\omega \in [0.1, 500]$ Hz

maximize gain : Level 1/2 = 6 [dB], Level 2/3 = 4 [dB]

maximize phase: Level 1/2 = 45 [deg], Level 2/3 = 35 [deg]

3. Stability margin of the rate feedback loop must be insured by limiting the upper bound of the closed-loop gain.

RaGMspc code:

closed loop, input 2 gyro noise input, output 2 closed-loop rate gain

soft constraint, $\omega \in [10^0, 10^3]$ Hz

minimize, Level 1/2 = [0 1.4], Level 2/3 = [-0.5 1.8]

4. Stability margin of the acceleration feedback loop must be insured by limiting the upper bound of the closed-loop gain.

AcGMspc code:

closed loop, input 1 disturbance torque, output 3 closed-loop acceleration gain

soft constraint, $\omega \in [10, 10^4]$ Hz

minimize, Level 1/2 = [0 1.4], Level 2/3 = [-0.5 1.8]

5. Two constraints are also included which are related to the simulation and not system performance itself. These involve the form of the two second-order compensators, one within the rate loop and the other within the acceleration loop.

The intent of the first such constraint is to insure a lead-lag form for the rate loop compensator. This requires the constraint that $a1 > 1$.

RComspc code:

soft constraint

maximize $dpp_6 = a1^2$, Level 1/2 = 1.01, Level 2/3 = 1

6. The intent of the acceleration loop constraint is to compensate for the structural anti-resonant-resonant characteristics of the gimbal. As such, the second-order filter must be of a lag-lead form which is insured by applying the constraint $\frac{1}{a2} > 1$. For optimization it is better to use a smooth performance criterion. After several trials it was determined that, instead of using $\frac{1}{a2} > 1$, it is better to use $0 < a2^2 < 1$, or $0 < dpp_11 < 1$. Denoting dp_11 as dpp_11 , the parameter is maintained positive. The upper limit is required by this spec.

Acmlspc code:

soft constraint

minimize *dpp*_11, Level 1/2 = 0.99, Level 2/3 = 1

3.4 Setup of the Initialization File

The performance chart of this problem does not contain time analysis, only frequency analysis. Thus linear simulation will not be performed and there is no need to specify *dt*. This problem is too fast (the fastest mode is 796 Hz for the nominal starting point) to be practically simulated in time.

epsilonF was set at 0.0002 (the default value) and 0 (machine precision), and tests were made for both values.

3.5 Optimization

For the antenna problem two starting points were considered. An acceptable starting point, called the nominal point, and an unacceptable starting point,

called the trivial point. The nominal point is defined by:

$$\begin{aligned}
 k1 &= 108 \frac{\text{v}}{\text{v}} \\
 f1 &= 8.6 \text{ Hz} \\
 f2 &= 417 \text{ Hz} \\
 f3 &= 250 \text{ Hz} \\
 z1 &= 0.707 \text{ Hz} \\
 a1 &= 2 \\
 k2 &= 500 \frac{\text{v}}{\text{v}} \\
 f4 &= 796 \text{ Hz} \\
 f5 &= 120 \text{ Hz} \\
 z2 &= 0.1 \\
 a2 &= 0.817 \text{ Hz}
 \end{aligned}$$

or equivalently approximately:

$$\begin{aligned}
 dpp_1 &= 1.08000e + 02 & dpp_7 &= 5.00000e + 02 \\
 dpp_2 &= 5.40300e + 01 & dpp_8 &= 5.00132e + 03 \\
 dpp_3 &= 2.62000e + 03 & dpp_9 &= 7.54000e + 02 \\
 dpp_4 &= 1.57100e + 03 & dpp_10 &= 2.00000e - 01 \\
 dpp_5 &= 1.41370e + 00 & dpp_11 &= 6.67500e - 01 \\
 dpp_6 &= 4.00000e + 00 & &
 \end{aligned}$$

and the trivial point by:

$$\begin{aligned}
k1 &= 50 \frac{\text{v}}{\text{v}} \\
f1 &= 5.0 \text{ Hz} \\
f2 &= 100 \text{ Hz} \\
f3 &= 100 \text{ Hz} \\
z1 &= 0.50 \text{ Hz} \\
a1 &= 1.5 \\
k2 &= 200 \frac{\text{v}}{\text{v}} \\
f4 &= 200 \text{ Hz} \\
f5 &= 100 \text{ Hz} \\
z2 &= 0.50 \\
a2 &= 0.667 \text{ Hz}
\end{aligned}$$

or equivalently approximately:

$$\begin{aligned}
dpp_1 &= 5.00000e + 01 & dpp_7 &= 2.00000e + 02 \\
dpp_2 &= 3.14150e + 01 & dpp_8 &= 1.25660e + 03 \\
dpp_3 &= 6.28300e + 02 & dpp_9 &= 6.28300e + 02 \\
dpp_4 &= 6.28300e + 02 & dpp_10 &= 1.00000e + 00 \\
dpp_5 &= 1.00000e + 00 & dpp_11 &= 4.44900e - 01 \\
dpp_6 &= 2.25000e + 00 & &
\end{aligned}$$

The optimization results are in Table 3.1. The optimal point achieved when the

Table 3.1: Comparison between the optimal points obtained by CONDUIT for the antenna problem

	Starting point	Optimal using epsilonF=0	Optimal using epsilonF=0.0002
Nominal	<i>MAX_COST_SOFT</i> 0.201512	<i>MAX_COST_SOFT</i> 0.0132284	<i>MAX_COST_SOFT</i> 0.0138598
Trivial	<i>MAX_HARD</i> 1.42554	<i>MAX_COST_SOFT</i> 2.26952	<i>MAX_COST_SOFT</i> 1.9644

optimization started in the nominal point and used epsilonF = 0 was:

$$\begin{aligned}
 dpp_1 &= 1.08002e + 02 & dpp_7 &= 5.00002e + 02 \\
 dpp_2 &= 5.40251e + 01 & dpp_8 &= 5.00132e + 03 \\
 dpp_3 &= 2.62000e + 03 & dpp_9 &= 7.54000e + 02 \\
 dpp_4 &= 1.57100e + 03 & dpp_10 &= 6.23329e - 02 \\
 dpp_5 &= 1.32943e + 00 & dpp_11 &= 6.20870e - 01 \\
 dpp_6 &= 3.98639e + 00
 \end{aligned}$$

The optimal point achieved when the optimization started in the nominal point and used epsilonF = 0.0002 was:

$$\begin{aligned}
 dpp_1 &= 1.08027e + 02 & dpp_7 &= 5.00014e + 02 \\
 dpp_2 &= 5.39746e + 01 & dpp_8 &= 5.00132e + 03 \\
 dpp_3 &= 2.62000e + 03 & dpp_9 &= 7.54001e + 02 \\
 dpp_4 &= 1.57100e + 03 & dpp_10 &= 4.88735e - 02 \\
 dpp_5 &= 1.30668e + 00 & dpp_11 &= 6.83348e - 01 \\
 dpp_6 &= 4.06439e + 00
 \end{aligned}$$

The largest percentage variation between these two optimal points is 21% for dp_10, and the largest variation is 0.0780 for dp_6.

The optimal point achieved when the optimization started in the trivial point and used $\epsilon_F = 0$ was:

$$\begin{aligned}
 dpp_1 &= 5.02402e + 01 & dpp_7 &= 2.00232e + 02 \\
 dpp_2 &= 3.07500e + 01 & dpp_8 &= 1.25660e + 03 \\
 dpp_3 &= 6.28322e + 02 & dpp_9 &= 6.28299e + 02 \\
 dpp_4 &= 6.28306e + 02 & dpp_{10} &= 4.65478e - 01 \\
 dpp_5 &= 6.93597e - 01 & dpp_{11} &= 1.01263e + 00 \\
 dpp_6 &= 5.25785e + 00
 \end{aligned}$$

The optimal point achieved when the optimization started in the trivial point and used $\epsilon_F = 0.0002$ was:

$$\begin{aligned}
 dpp_1 &= 5.18130e + 01 & dpp_7 &= 2.02086e + 02 \\
 dpp_2 &= 2.57491e + 01 & dpp_8 &= 1.25658e + 03 \\
 dpp_3 &= 6.28449e + 02 & dpp_9 &= 6.28298e + 02 \\
 dpp_4 &= 6.28578e + 02 & dpp_{10} &= 7.90417e - 01 \\
 dpp_5 &= 5.85324e - 01 & dpp_{11} &= 1.00956e + 00 \\
 dpp_6 &= 4.92225e + 00
 \end{aligned}$$

The largest percentage variation between these two optimal points is 69% for dp_10, and the largest variation is 5 for dp_2.

It can be seen in Table 3.1 that for this problem the performance is better for the optimization using $\epsilon_F=0.0002$.

The performance for the problem starting in the nominal point and optimized with $\epsilon_F=0.0002$ is presented in Figure 3.2 and Figure 3.3. By inspecting Figure 3.2 it can be seen that there is only one spec, AcGMspc, which does not

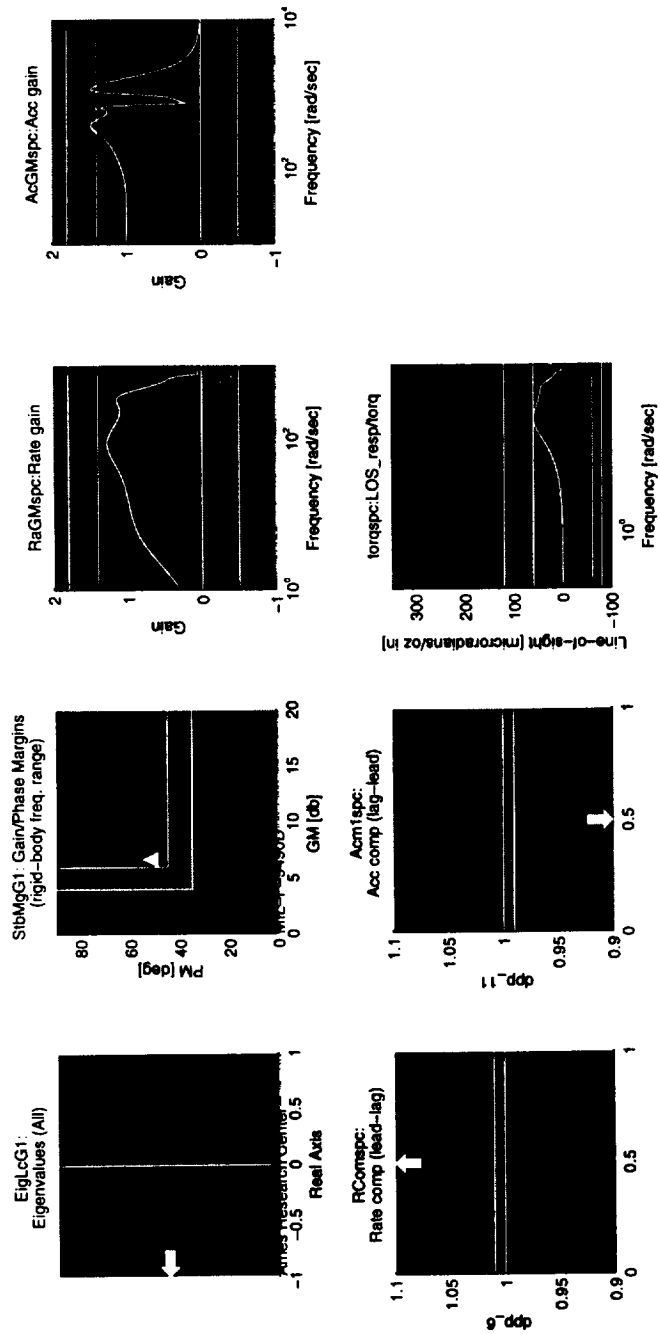


Figure 3.2: Antenna problem, nominal starting point

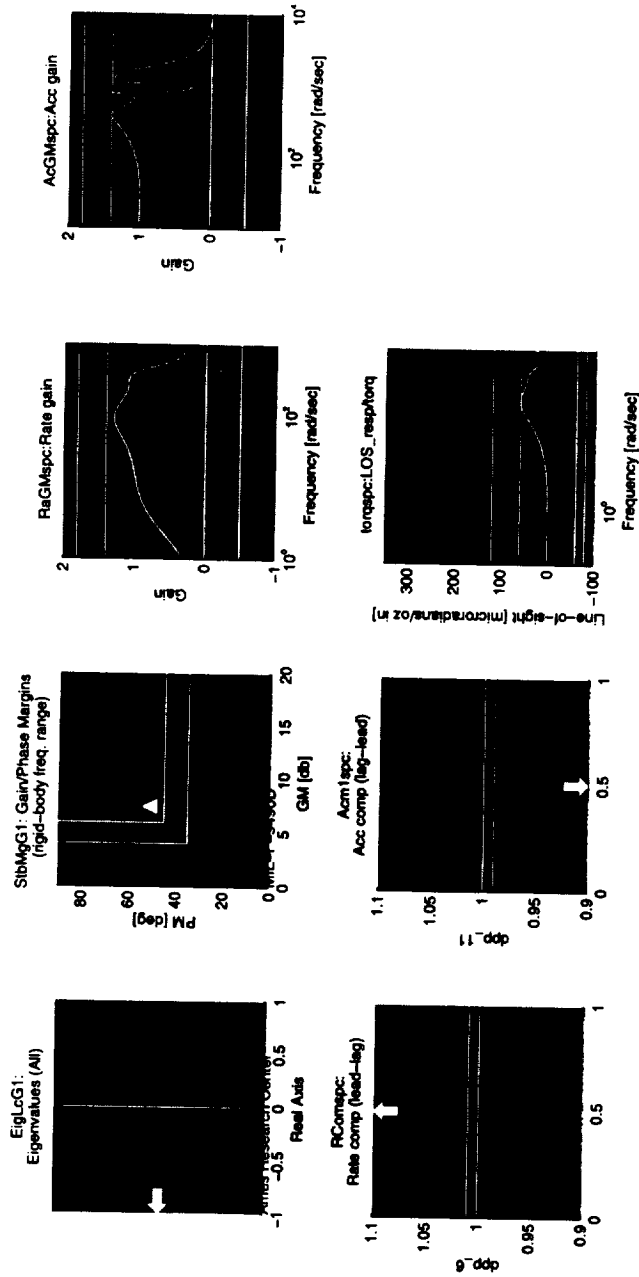


Figure 3.3: Antenna problem, nominal point improved by CONDUIT, $\epsilon_F=0.0002$

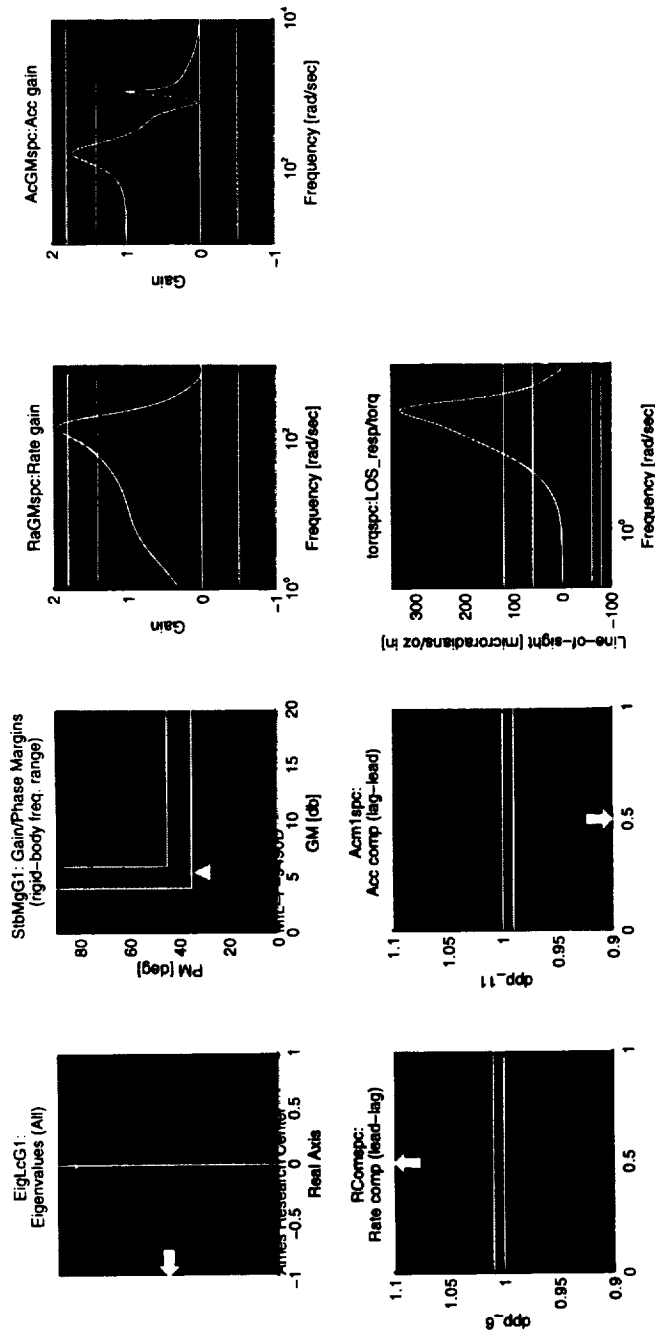


Figure 3.4: Antenna problem, trivial starting point

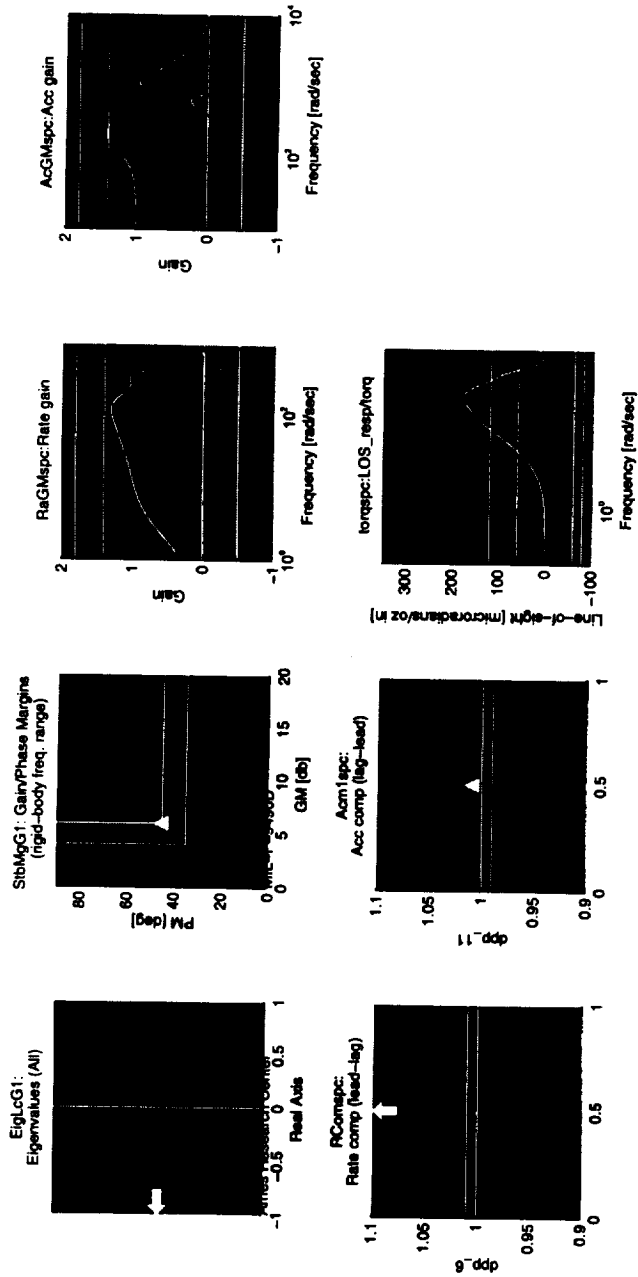


Figure 3.5: Antenna problem, trivial point improved by CONDUIT, $\epsilon_F=0.0002$

satisfy Level 1 at the starting point and it is improved through optimization, Figure 3.3.

The performance for the problem starting in the trivial point and optimized with $\epsilon_F=0.0002$ is presented in the Figure 3.4 and Figure 3.5. By inspecting Figure 3.4 and Figure 3.5 it can be seen that the stability margins improved from the Level 3 (red zone) to the Level 1 (blue zone). The improved point has gain margin=6 and phase margin=45.5, being near the border of Level 1/2. The gain in stability was obtained at the expense of a soft constraint, $Acmlspc$, that couldn't be satisfied. Optimization changed the unacceptable initial point to an acceptable optimal point. The unsatisfied spec, $Acmlspc$, is not important, because it is an artificial spec that requires the compensator in the acceleration loop to be lag-lead. Optimization showed that for these parameters a lead-lag works better.

3.6 Optimization Attempts

As a cautionary note, one has to be very careful with the constraints.

Originally the antenna problem didn't contain the $EigLcG1$ and $StbMgG1$ specs, and was optimized using a simplified performance chart containing only five specifications: $RaGMspc$, $AcGMspc$, $RComspc.m$, $Acmlspc$, $torqspc$. Figure 3.6 and Figure 3.7 show the results of this optimization attempt. It can be seen that even though the system is stable at the starting point, it is driven to instability. This is obviously unsatisfactory, regardless of the values of the specifications actually included in the problem formulation.

In addition to introducing the stability specs, another improvement done for

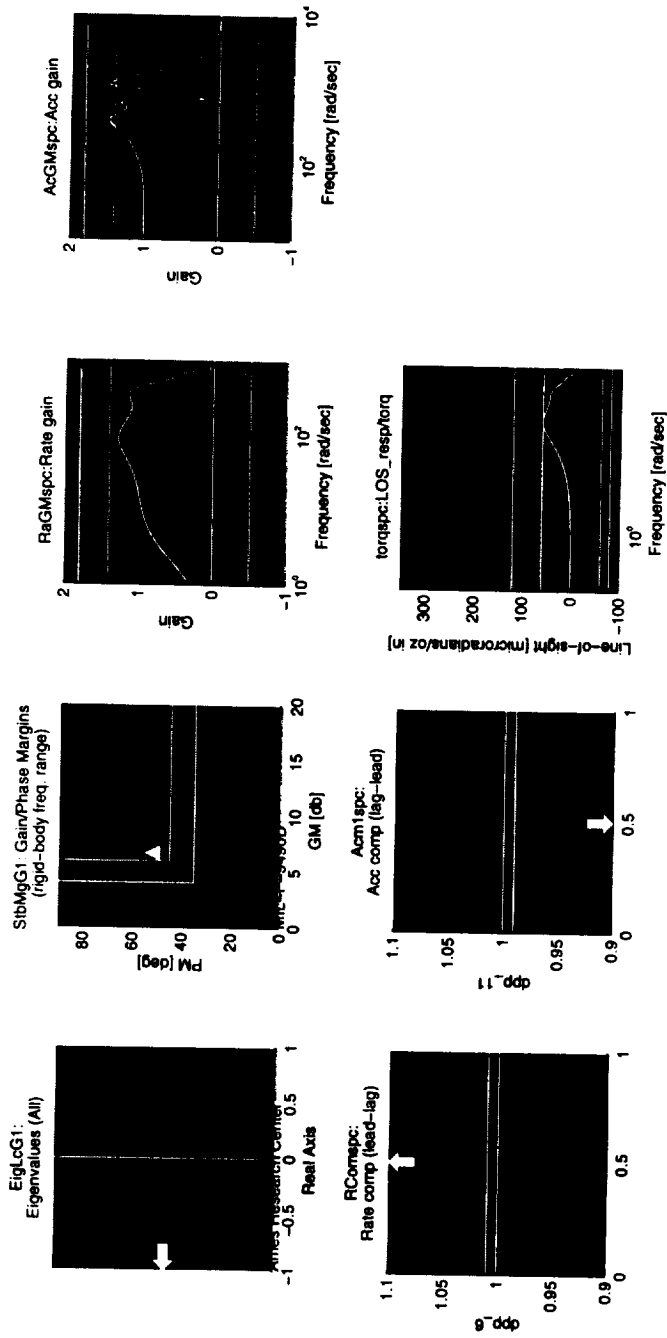


Figure 3.6: Antenna problem, nominal starting point

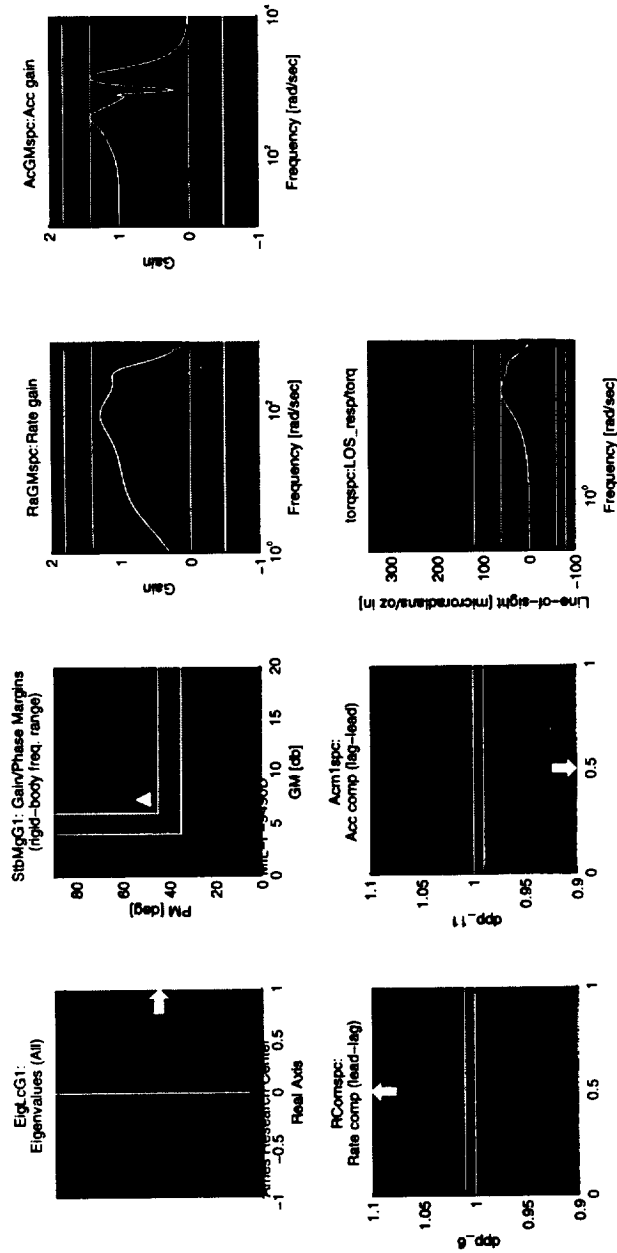


Figure 3.7: Antenna problem, nominal point is driven unstable when stability specs are check only, epsilonF=0.0002

this problem was the restriction that all design parameters be positive. Otherwise they can become negative, resulting in unstable or nonminimum phase compensators.

Chapter 4

Sensitivity

CONDUIT provides an optimal set of design parameters, $\hat{\theta}$, with respect to the given specifications. But the remaining question is if the controller model proposed by the designer is good enough. In a controller design problem, the sensitivity issue occurs in several ways:

1. performance sensitivity to plant parameter changes
2. performance sensitivity to controller parameter changes
3. parameter sensitivity to small changes in the constraints

An analogous sensitivity question arises in maximum likelihood estimation (MLE). Effective tools for analyzing sensitivity in the MLE situation are well known [16]. These tools have been implemented in a software package for aircraft identification and extensively tested and used [23].

The main focus of the work described here was on the extension of the MLE ideas and methods to CONDUIT and the control design problem.

4.1 The Different Forms of Sensitivity

4.1.1 Performance sensitivity to plant parameter changes

Changes in the plant parameters are very much like disturbance signals and are generally addressed by insisting on robust controllers. Thus, the sensitivity of the design to changes in plant parameters is covered by the specifications, especially those relating to stability margins.

4.1.2 Performance sensitivity to controller parameter changes

Min max optimization is trying to minimize the maximum objective (constraint) at every iteration. This function is called MAX_ACTIVE in C-O. The MAX_ACTIVE function, called MAX_HARD (see section 1.2) in phase 1, MAX_COST_SOFT in phase 2, or MAX_COST in phase 3, is the function that is actually being minimized. The possible situations when optimization stops are presented in Figures 4.1–4.8. Note that, in these figures, specifications are indicated by thin lines while the active spec (MAX_ACTIVE) is indicated by a heavy line.

- One or some hard constraints cannot be satisfied, and the optimization stops in Phase 1, Figure 4.1 and Figure 4.2. For these cases, the sensitivity computation is unnecessary, because a feasible set of parameters and an optimal point do not exist.
- All hard constraints are satisfied, one or more soft constraints cannot be satisfied, and the optimization stops in Phase 2, Figures 4.3-4.5.

MAX_ACTIVE

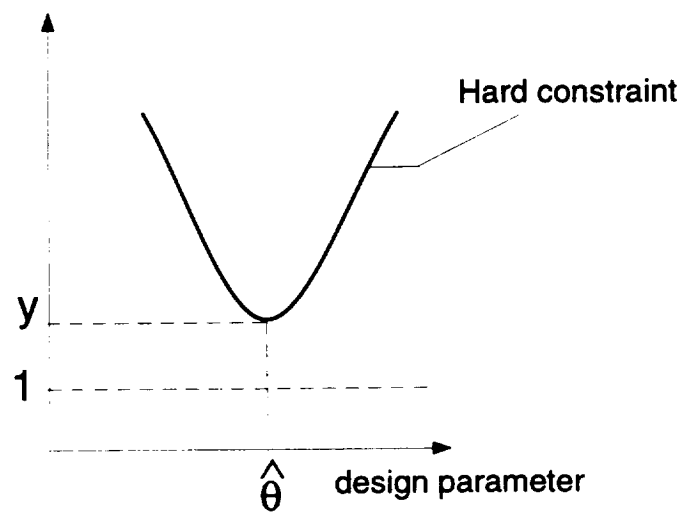


Figure 4.1: The minimum is Phase 1

MAX_ACTIVE

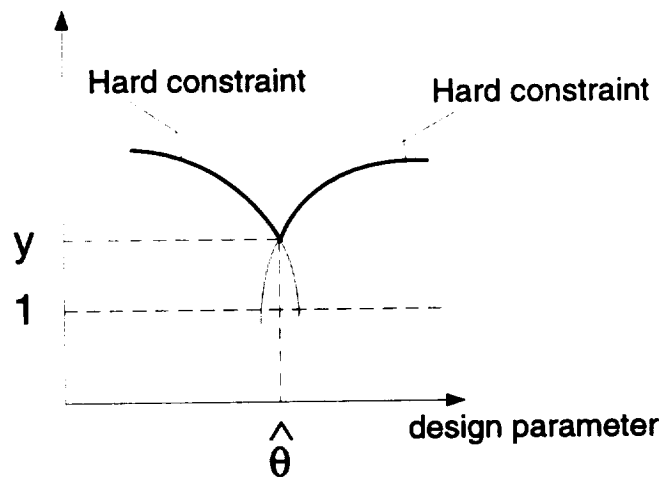


Figure 4.2: The minimum is Phase 1

It is also possible that the optimization stops in Phase 2 when an objective competes with a soft constraint. But in this situation the optimization problem is not well formulated and it is better to change that objective so it is satisfied. This way the objective will allow the soft constraint to be improved first, and then, in Phase 3, the objective will be improved to its best value.

In Figure 4.5 it can be seen that in the region where the

$$\text{hard constraint} > 1, \quad (4.1)$$

MAX_ACTIVE equals the hard constraint even if there is a larger soft constraint.

When the minimum is in Phase 2 the optimization is considered acceptable.

- When all hard and soft constraints are met, the optimization stops in Phase 3, as shown in Figures 4.6, 4.7, and 4.8.

Figure 4.8 shows a situation where the constraints are met (< 1), in the region surrounding $\hat{\theta}$; MAX_ACTIVE equals the objective even if there is a constraint with a larger value. This case, in which the optimization pushes a constraint to its border by minimizing an objective is the most common case.

When the minimum is in Phase 3 the optimization is considered satisfactory.

MAX_ACTIVE

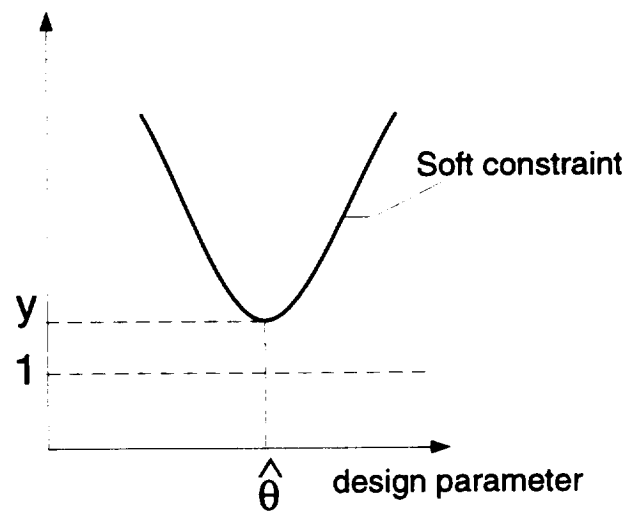


Figure 4.3: The minimum is Phase 2

MAX_ACTIVE

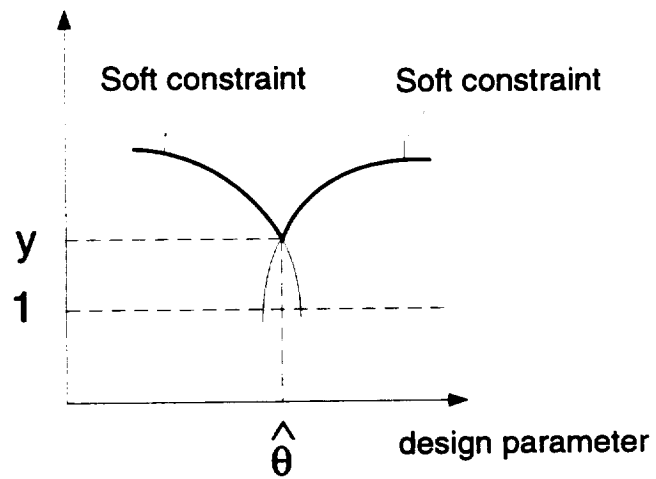


Figure 4.4: The minimum is Phase 2

MAX_ACTIVE

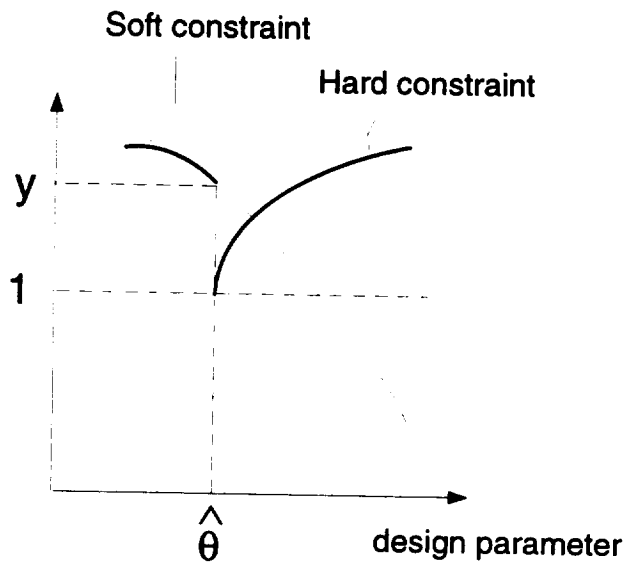


Figure 4.5: The minimum is Phase 2

MAX_ACTIVE

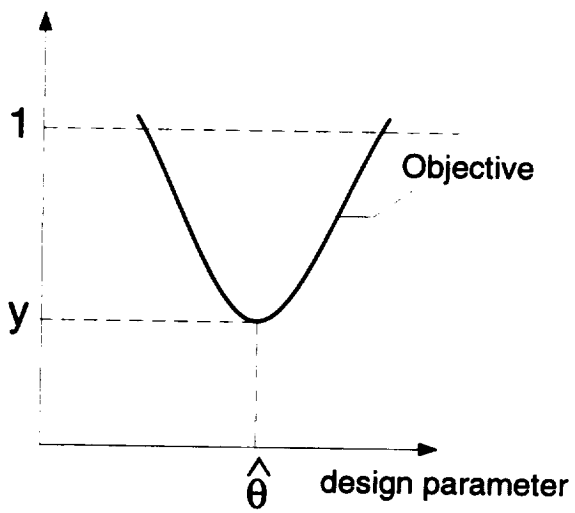


Figure 4.6: The minimum is Phase 3

MAX_ACTIVE

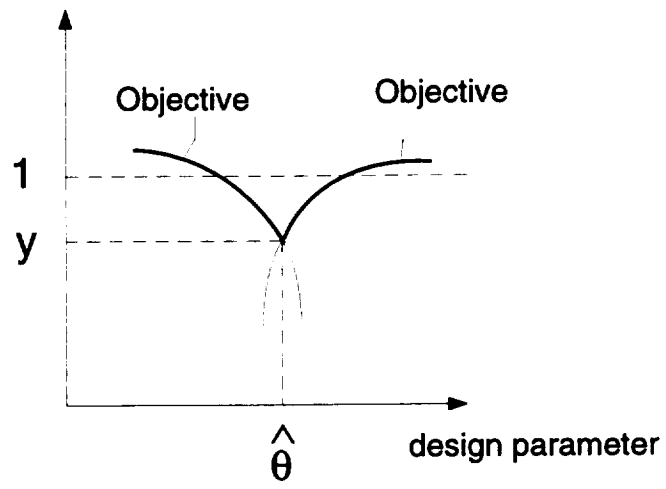


Figure 4.7: The minimum is Phase 3

MAX_ACTIVE

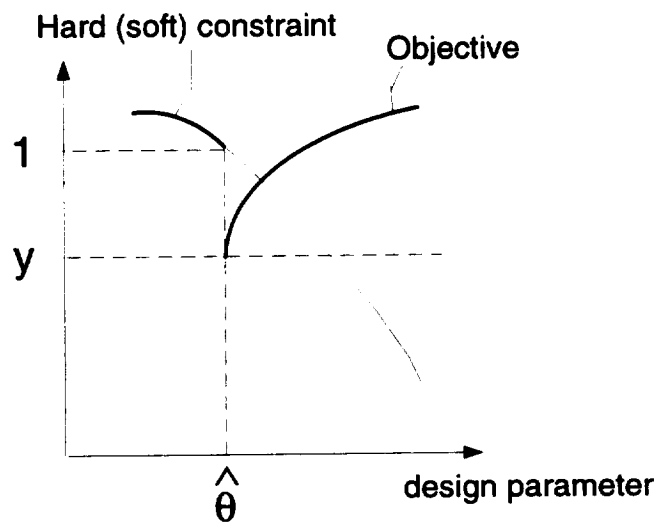


Figure 4.8: The minimum is Phase 3

These minimum points can be characterized by considering the local geometry in two classes:

A) When the minimum occurs at a point similar to those shown in Figures 4.3 and 4.6, it is equivalent to the solution to an unconstrained optimization problem. MAX_ACTIVE has:

- Gradient zero on all design parameter axes at the optimal point
- Hessian positive semidefinite at the optimal point

and J, the function used in sensitivity computation, equals MAX_ACTIVE.

B) When the minimum looks as in Figures 4.4, 4.5, 4.7, and 4.8 MAX_ACTIVE has:

- Gradient discontinuous on at least one design parameter axis
- Hessian undefined.

The sensitivity test implementation includes the gradient computation of all optimization specs and of J, at the optimal point. In case the gradient of J is zero, then the optimal point is case A) and insensitivities, correlations, and Cramér-Rao bounds can be computed as in section 4.2. In case the gradient of J is discontinuous, then section 4.3 applies.

4.1.3 Parameter sensitivity to small changes in the constraints

The sensitivity of the optimal solution to small modifications in the constraints is an important aspect of practical applications. The sensitivity theorem will be presented in section 4.4.

4.2 Sensitivity Tools for Use When the Gradient of MAX_ACTIVE is zero at the optimal point

4.2.1 Insensitivity Ellipsoid

- **Estimation**

In estimation theory, the confidence region of an estimator is often evaluated using the probability density function of the estimate, $p_{\hat{\theta}}(\theta)$, i.e. the probability that the estimated $\hat{\theta}$ is actually θ . If the maximum likelihood estimation technique is used to estimate a parameter vector θ , this probability density function is approximately Gaussian, unbiased, and with covariance matrix equal to $M(\theta_t)^{-1}$:

$$p_{\hat{\theta}}(\theta) \cong [(2\pi)^n |M(\theta_t)^{-1}|]^{-1/2} \exp\left\{-\frac{1}{2}(\theta - \theta_t)^T M(\theta_t)(\theta - \theta_t)\right\} \quad (4.2)$$

where M is the Fischer information matrix:

$$M(\theta_t) \equiv E\{[\nabla_{\theta_t} \ln p(Z|\theta_t)][\nabla_{\theta_t}^T \ln p(Z|\theta_t)|\theta_t]\} \cong -E\{\nabla_{\theta_t}^2 \ln p(Z|\theta_t)|\theta_t\} \quad (4.3)$$

where Z represents the observations and θ_t denotes the true value of θ [17].

The equation

$$(\theta - \theta_t)^T M(\theta_t)(\theta - \theta_t) = Constant \quad (4.4)$$

describes isometric surfaces (ellipsoids) centered at θ_t , with the same value for the probability density. For *Constant* equal one, the volume integral over the interior of the ellipsoid is 63.2% and is equal to the probability that $\hat{\theta}$ lies within the ellipsoid. From the practical point of view θ_t is not

available and the above ellipsoid cannot be drawn. This problem is handled by reversing the roles of θ_t and $\hat{\theta}$, because $M(\hat{\theta})$ can approximate $M(\theta_t)$. The ellipsoid defined by

$$(\theta - \hat{\theta})^T M(\hat{\theta})(\theta - \hat{\theta}) = 1 \quad (4.5)$$

is called the **uncertainty ellipsoid**. For maximum likelihood estimation, if the negative log likelihood function J is:

$$J(\theta) = \frac{1}{2} \sum_{i=1}^N [\tilde{z}_\theta(t_i) - z(t_i)]^T R^{-1} [\tilde{z}_\theta(t_i) - z(t_i)] \quad (4.6)$$

the information matrix can be closely approximated by

$$H \equiv \sum_{i=1}^N \nabla_\theta \tilde{z}_\theta^T(t_i) R^{-1} \nabla_\theta^T \tilde{z}_\theta(t_i) \quad (4.7)$$

z is the observation vector, \tilde{z} is the predicted value of z if $\theta = \hat{\theta}$, R is the covariance matrix, N is the number of measurements of z . The H-matrix is identical to the dominant term of the Hessian matrix $\nabla_\theta^2 J(\theta)$.

• Optimization

The estimation problem is to find the model parameters, $\hat{\theta}$, that minimize the negative log likelihood function J . This function expresses the difference between the predicted output based on $\hat{\theta}$ and the measured output.

The optimization problem is to find the controller parameters, $\hat{\theta}$, that minimize the MAX_ACTIVE function, denoted also by J . This function expresses the difference between the predicted performance based on $\hat{\theta}$ and the desired performance. MAX_ACTIVE is called MAX_HARD in Phase 1, MAX_COST_SOFT in Phase 2, or MAX_COST in Phase 3 of the optimization. Because there is no probability in the optimization case

$$M(\hat{\theta}) = \nabla_\theta^2 J(\hat{\theta}). \quad (4.8)$$

For case A) MAX_ACTIVE can be expanded about θ_t as follows:

$$J(\theta) = J(\theta_t) + \frac{1}{2}(\theta - \theta_t)^T \nabla_{\theta}^2 J(\theta_t)(\theta - \theta_t) + O((\theta - \theta_t)^3) \quad (4.9)$$

θ_t represents the parameters that achieve the best performance and $\hat{\theta}$ the optimal parameters as determined computationally. Suppose that $J(\hat{\theta}) - J(\theta_t)$ is less than some constant. This situation is referred to as a sub-optimal solution. This can be sufficient where the exact optimum is not required, but performance is required to be within a specified range of the optimum. If higher order terms are ignored, this condition becomes

$\frac{1}{2}(\hat{\theta} - \theta_t)^T \nabla_{\theta}^2 J(\theta_t)(\hat{\theta} - \theta_t) \leq \text{Constant}$ and can be approximately as before with

$$\frac{1}{2}(\theta_t - \hat{\theta})^T \nabla_{\theta}^2 J(\hat{\theta})(\theta_t - \hat{\theta}) \leq \text{Constant} \quad (4.10)$$

This ellipsoid is essentially the same as the uncertainty ellipsoid if the constant is chosen to be $\frac{1}{2}$. But because it is described using MAX_ACTIVE and not the negative log likelihood, we named it the **insensitivity ellipsoid**.

In the CONDUIT optimization problem the fit error, $\nabla_{\theta}^2 J(\theta)$, often cannot be approximated by H. If MAX_ACTIVE were computed as the sum of errors of performance functions then the Hessian of MAX_ACTIVE could be approximated by the sum of square gradients of all the optimization specs. But MAX_ACTIVE is computed as the max error of performance functions and shows the behavior of only one performance function at a time, without any influence of the others. Thus the Hessian of MAX_ACTIVE generally shows the curvature of only one performance function at a time. Using the squared gradient of MAX_ACTIVE to approximate the Hessian

frequently results in a singular matrix so the estimation theory approximation cannot be applied. These limitations do not apply if an objective is formed by summing a collection of specs. It is recommended that the summed objective be used in CONDUIT.

Moreover, if M is approximated with the sum of squared gradients or the sum of Hessians of all optimization specs, then there is not certitude that M is positive semidefinite. The specs are randomly convex, concave or neither at the optimal point. But the Hessian of J is positive semidefinite, as will be shown in section 4.2.5.

The insensitivity ellipsoid can be thought of as the result of minimizing the MAX_ACTIVE. This minimization is a form of a sensitivity analysis based on the principle that two values of θ cannot be reliably distinguished unless they result in sufficient differences in the MAX_ACTIVE function.

Figure 4.9 shows the insensitivity ellipsoid in the case of two unknowns. The interior of the insensitivity ellipse is shaded. Most of the concepts discussed below can be illustrated by using such a two-dimensional ellipse.

Optimization problems solved with CONDUIT have typically more than 10 parameters; that is, the ellipsoid is more than 10 dimensional. Because it is hard to visualize the information on the ellipsoid, there are three other formats for displaying it: insensitivity, correlation, and Cramér-Rao bound [23].

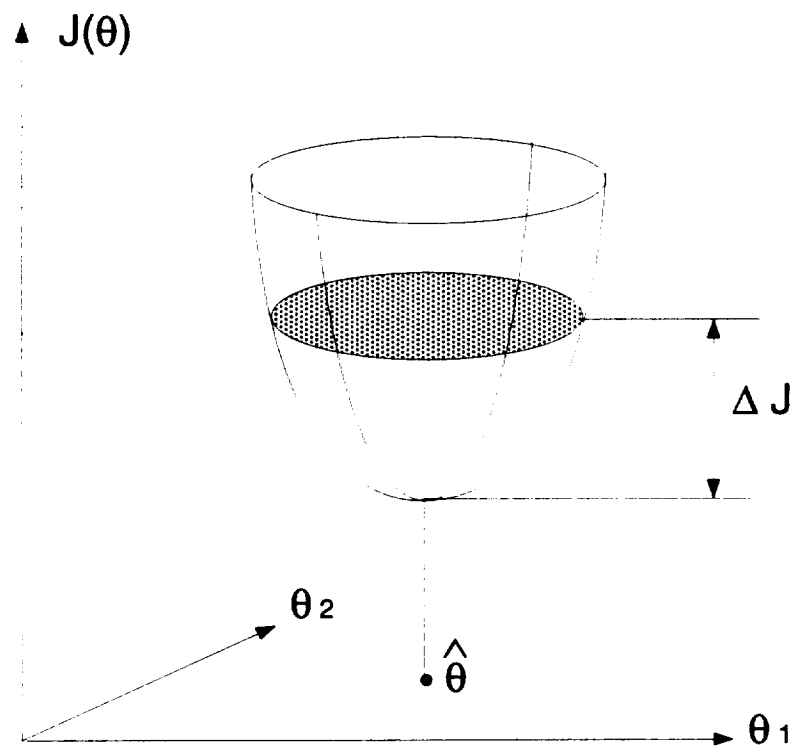


Figure 4.9: Two-dimensional insensitivity ellipsoid

4.2.2 Insensitivity

Insensitivity is a measure of how much one parameter can be changed from the achieved optimal value without causing the performance to degrade by more than a given amount. To simplify the computation, J is approximated by its Hessian (second order approximation), M , at the optimal point. This way, instead of recomputing J for parameter changes, only two more evaluations in addition to the optimal point evaluation are needed.

An expression for the insensitivity can be obtained from the equation for the insensitivity ellipsoid, Equation 4.10. Investigate changes in $\hat{\theta}$ along the direction of the i th parameter, that is $\theta - \hat{\theta} = \alpha e_i$, where e_i is a unit vector in the direction of the i th parameter and α is an arbitrary scalar. To stay inside the ellipsoid, the following must be true:

$$\alpha^2 e_i^T M e_i \leq 1 \quad (4.11)$$

Since $e_i^T M e_i = M_{ii} \geq 0$ because M is positive semidefinite (see section 4.2.5), then

$$|\alpha| \leq (M_{ii})^{-1/2} \quad (4.12)$$

The maximum magnitude of α is given by the equality, which occurs at the intersection of the e_i axis with the ellipsoid. This value is defined as the insensitivity with respect to θ_i . Figure 4.10 illustrates the geometric interpretation of the insensitivity for a problem with two unknowns. For convenience, the origin has been redefined to be the center of the ellipse, which is the maximum likelihood estimate. The insensitivities of θ_1 and θ_2 are labeled as I_1 and I_2 .

Remark 1. The insensitivity is a reasonable measure of accuracy only when a single parameter is being estimated, because it ignores any effect of correlation between parameters. When several parameters are estimated, the error band is

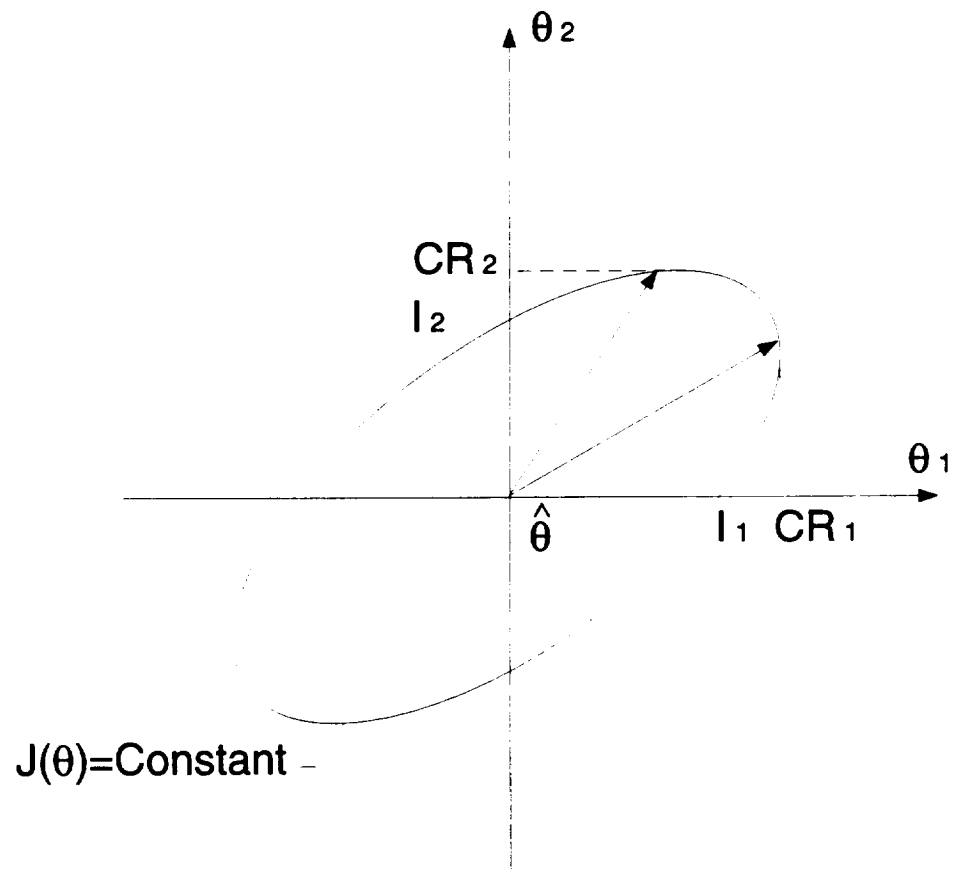


Figure 4.10: Geometric interpretation of insensitivities and Cramér-Rao bounds

a result of both insensitivity and correlations. Insensitivity gives only the lower bound for the error band, but correlation effects tend to increase the error band so much that the insensitivity is virtually useless as an indicator of accuracy.

Remark 2. In order to compare insensitivities of different parameters to each other, they must be scaled with the parameter values $\frac{I_i}{\theta_i}$.

4.2.3 Cramér-Rao Bound

The Cramér-Rao bound is also based on the insensitivity ellipsoid [17]. The standard deviation of the i th parameter is $\sqrt{(M^{-1})_{ii}}$. For maximum likelihood estimates, the Cramér-Rao inequality gives the approximation H^{-1} for the covariance matrix and the quantity $\sqrt{(H^{-1})_{ii}}$ is called the Cramér-Rao bound. In this thesis J is different from maximum likelihood estimation and the Cramér-Rao bound will be defined to be equal to the standard deviation. The Cramér-Rao bounds and insensitivities are closely related. While the insensitivity is the conditional standard deviation of the parameter estimate, given that all the other parameters are known, the Cramér-Rao bound is the unconditional standard deviation. The insensitivity is $\sqrt{(M_{ii})^{-1}}$ and the Cramér-Rao bound is $\sqrt{(M^{-1})_{ii}}$. The geometric relationship between these quantities is the following. The insensitivity is the solution to the following optimization problem: maximize $|\theta_i - \hat{\theta}_i|$ while staying within the insensitivity ellipsoid and holding all of the other parameters fixed. The Cramér-Rao bound is the solution to the same problem without holding the other parameters fixed. The removal of the restrictions on the optimization is directly analogous to removing the statistical conditioning. A two-dimensional plot of the geometric relationship between the two measures is shown in Figure 4.10. The insensitivities of θ_1 and θ_2 are labeled I_1 and I_2 ,

while the Cramér-Rao bounds are labeled CR_1 and CR_2 .

The proof of the affirmation that the Cramér-Rao bound is the solution to the optimization problem above comes from the next theorem.

Theorem 4.1 *Given a fixed vector y and a positive definite symmetric matrix M , then the maximum of $x^T y$, subject to the constraint that $x^T M x \leq 1$, is given by $\sqrt{y^T M^{-1} y}$ and the maximum is obtained for $x = \frac{M^{-1} y}{\sqrt{y^T M^{-1} y}}$.*

Apply this theorem to the insensitivity ellipsoid, where y is e_i , a unit vector along the axis of the i th parameter, and x is $\theta - \hat{\theta}$. The maximum of $\theta_i - \hat{\theta}_i = (\theta - \hat{\theta})^T e_i$ is

$$\sqrt{e_i^T M^{-1} e_i} = \sqrt{(M^{-1})_{ii}} \quad (4.13)$$

and this is CR_i , the Cramér-Rao bound. CR_i is the largest projection on the θ_i axis of the vectors with origin in $\hat{\theta}$ and arrow on the ellipsoid. It is caused by the vector :

$$\theta - \hat{\theta} = \frac{M^{-1} e_i}{\sqrt{(M^{-1})_{ii}}} \quad (4.14)$$

This projection shows the largest error that can occur for θ_i without affecting performance. This largest error for θ_i is unperceived only if all the other parameters have particular values, so that together they form the vector $\theta - \hat{\theta} = \frac{M^{-1} e_i}{\sqrt{(M^{-1})_{ii}}}$.

Remark 1. In order to compare Cramér-Rao bounds of different parameters to each other, they must be scaled with the parameter values, $\frac{CR_i}{\theta_i}$.

Remark 2. To see the correlations that may cause a big Cramér-Rao bound, CR_i , the vector $\theta - \hat{\theta}$ that has this projection on θ_i must be decomposed in the vector space formed by all insensitivities vectors: $I_1 * e_1, \dots, I_n * e_n$, where e_i is a unit vector along the axis of the i th parameter. This is equivalent to have an

ellipsoid that cuts all axes at 1, all insensitivities being 1. Figure 4.11 shows the projections p_1, \dots, p_n .

$$\theta - \hat{\theta} = \frac{M^{-1}e_i}{\sqrt{(M^{-1})_{ii}}} = [p_1 \ p_2 \ \dots \ p_n]^T = [a_1 * I_1 \ a_2 * I_2 \ \dots \ a_n * I_n]^T \quad (4.15)$$

so that

$$[a_1 \ a_2 \ \dots \ a_n]^T = \Theta_{CR_i} = \frac{T^{-1}M^{-1}(:, i)}{CR_i} \quad (4.16)$$

where $M^{-1}(:, i)$ denotes the i th column of M^{-1} and $T = \text{diag}[I_1, I_2, \dots, I_n]$. The large components of Θ_{CR_i} show the parameters that are correlated with θ_i . In this case the table showing the correlations is not symmetrical.

We can compare the projections of the vector $\theta - \hat{\theta}$ in the vector space $CR_1 * e_1, \dots, CR_n * e_n$, where e_i is a unit vector along the axis of the i th parameter. This is equivalent to having an ellipsoid that can be enclosed in an n -dimensional hypercube with all sides 2, all CR_i being 1. Then the correlation table is symmetrical, with all diagonal elements equal to 1. Both scalings (with insensitivities and with Cramér-Rao bounds) are equally good, and the first one is presented in the example in subsection 4.2.8.

4.2.4 Eigenvalues and Eigenvectors

The parameter correlation can be seen also by studying the eigenvalues and eigenvectors of M , normalized to have unity diagonal elements. The unnormalized M matrix is of very little use in studying correlation, because scaling effects tend to dominate. The normalization is done by dividing every row i and column i by $\sqrt{M(i, i)}$. The resulting ellipsoid will intersect every axis at 1, having the insensitivity of all parameters equal 1.

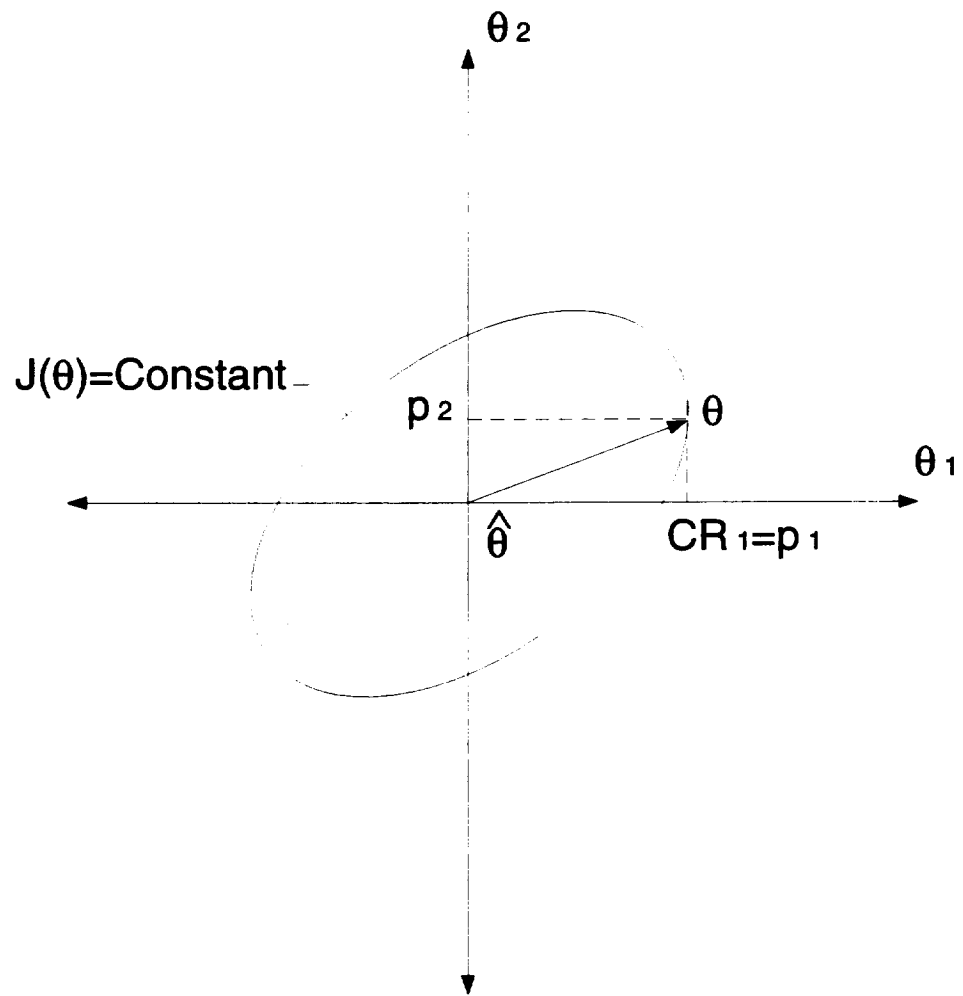


Figure 4.11: Projections

If $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of the normalized M matrix, then the diagonals of the ellipsoid are $2\frac{1}{\sqrt{\lambda_i}}$. Small eigenvalues of the normalized M matrix indicate correlations among the parameters with significant components in the corresponding eigenvectors.

Eigenvalues and eigenvectors values of the Hessian of MAX_ACTIVE are interpreted in the context of parameter correlation in an example at the end of subsection 4.2.8.

4.2.5 Positive Semidefiniteness

In order to compute the Cramér-Rao bounds, a positive definite M is necessary. The question is if the Hessian of MAX_ACTIVE is always positive. The answer is usually yes, but under certain circumstances it could also be zero.

As can be seen in Figure A), the Hessian is positive definite. It is also possible that MAX_ACTIVE reached a minimum in some direction, but does not change at all in some other directions.

If MAX_ACTIVE is completely insensitive in a direction that coincides with a parameter axis, then the insensitivity ellipsoid is degenerate on that parameter axis, Figure 4.12. Therefore the Hessian is singular, insensitivity is infinite for that parameter and the Cramér-Rao bounds are also infinite. This may happen because that parameter does not change locally, an acceptable behavior, or globally, in which case it should be eliminated.

If MAX_ACTIVE is completely insensitive in a direction that does not coincide with a parameter axis, then the insensitivity ellipsoid is degenerate in that direction, but still cuts the axes, Figure 4.13. Therefore the Hessian is also singular, insensitivities are finite, and the Cramér-Rao bounds are infinite. This

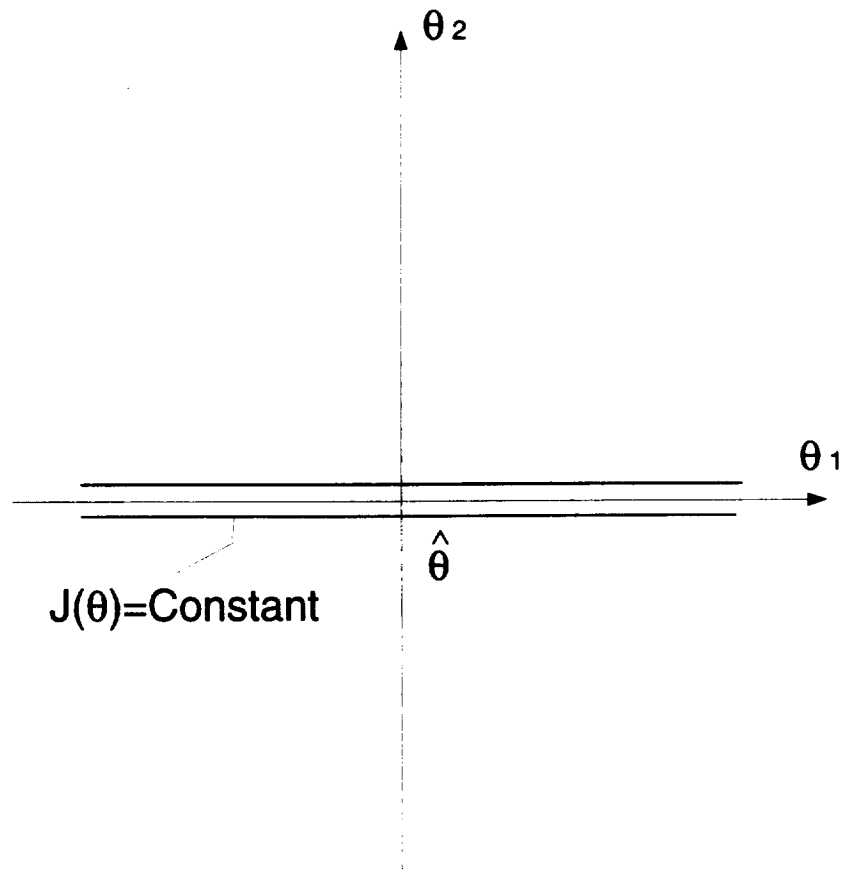
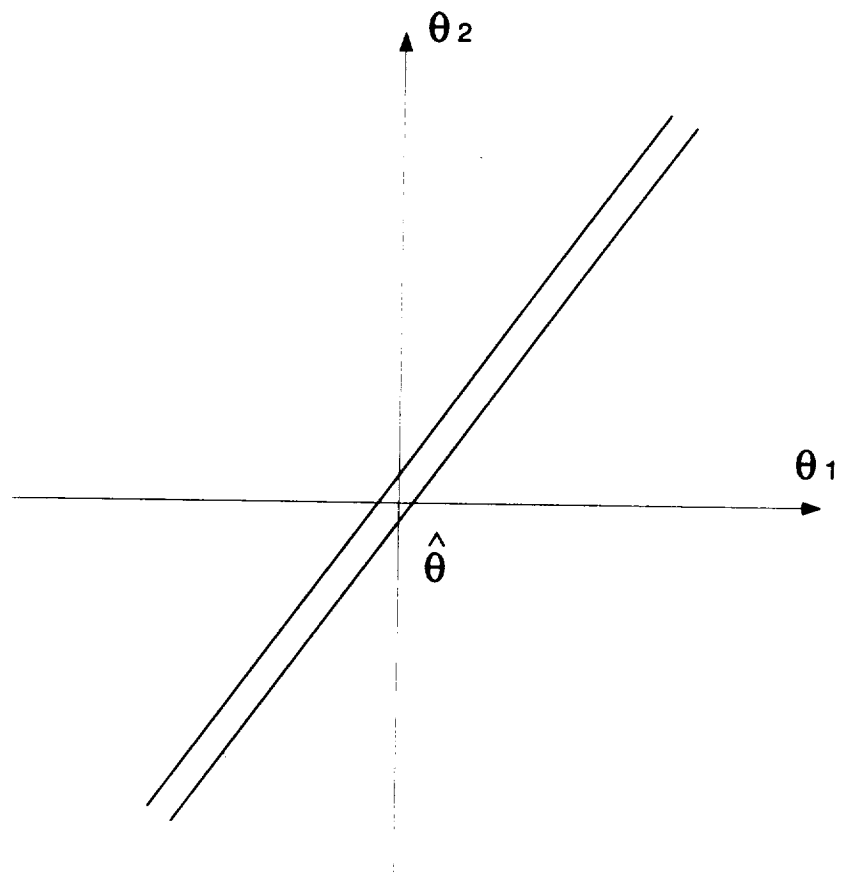


Figure 4.12: Ellipsoid is degenerate



$$J(\theta) = \text{Constant}$$

Figure 4.13: Ellipsoid is degenerate

shows correlation, and one of the parameters involved in the correlation should be eliminated.

4.2.6 Hessian Formulae

To compute the Hessian one can use, among others, the following three methods:

- Method one (faster):

$$\text{Hess}_{jk}(f, h) \approx \frac{1}{2} * \left\{ \frac{f(x_j + h, x_k + h) - 2f(x_j, x_k) + f(x_j - h, x_k - h)}{h^2} - \frac{f(x_j + h, x_k) - 2f(x_j, x_k) + f(x_j - h, x_k)}{h^2} - \frac{f(x_j, x_k + h) - 2f(x_j, x_k) + f(x_j, x_k - h)}{h^2} \right\} \quad (4.17)$$

For this method one needs in general $N^2 + N$ function evaluations, where N is the number of the parameters. First the diagonal terms are evaluated using $2 * N$ function evaluations and then the upper diagonal elements needs only $N^2 - N$ function evaluations.

The truncation error for off-diagonal elements is

$$\frac{h^2}{12} \left[2 \frac{\partial^4 f}{\partial x_j^3 \partial x_k} (x_j, x_k) + 3 \frac{\partial^4 f}{\partial x_j^2 \partial x_k^2} (x_j, x_k) + 2 \frac{\partial^4 f}{\partial x_j \partial x_k^3} (x_j, x_k) \right] + \text{Higher order terms} \quad (4.18)$$

The truncation error for diagonal elements is

$$\frac{h^2}{12} \left[\frac{\partial^4 f}{\partial x_j^4} \right] + \text{Higher order terms} \quad (4.19)$$

The roundoff error for off-diagonal elements is

$$\frac{12\epsilon_f f}{2h^2} \quad (4.20)$$

The roundoff error for diagonal elements is

$$\frac{4\epsilon_f f}{h^2} \quad (4.21)$$

- Method two (slower):

$$\text{Hess}_{jk}(f, h) \approx \left[\frac{f(x_j + h, x_k + h) - f(x_j + h, x_k - h)}{4h^2} \right] - \left[\frac{f(x_j - h, x_k + h) - f(x_j - h, x_k - h)}{4h^2} \right] \quad (4.22)$$

For this method one needs in general $2 * N^2$ function evaluations, where N is the number of the parameters.

The truncation error for off-diagonal elements is

$$\frac{h^2}{6} \left[\frac{\partial^4 f}{\partial x_j^3 \partial x_k}(x_j, x_k) + \frac{\partial^4 f}{\partial x_j \partial x_k^3}(x_j, x_k) \right] + \text{Higher order terms} \quad (4.23)$$

The truncation error for diagonal elements is

$$\frac{h^2}{12} \left[\frac{\partial^4 f}{\partial x_j^4} \right] + \text{Higher order terms} \quad (4.24)$$

The roundoff error for off-diagonal elements is

$$\frac{4\epsilon_f f}{4h^2} \quad (4.25)$$

The roundoff error for diagonal elements is

$$\frac{4\epsilon_f f}{h^2} \quad (4.26)$$

- Method three:

$$\text{Hess}_{jk}(f, h_j, h_k) \approx \left[\frac{f(x_j + h_j, x_k + h_k) - f(x_j + h_j, x_k - h_k)}{4h_j h_k} \right] - \left[\frac{f(x_j - h_j, x_k + h_k) - f(x_j - h_j, x_k - h_k)}{4h_j h_k} \right] \quad (4.27)$$

where $h_j = h * \max(1, \text{abs}(x_j))$.

For this method one needs in general $2 * N^2$ function evaluations, where N is the number of the parameters.

The truncation error is

$$\left[\frac{h_j^2}{6} \frac{\partial^4 f}{\partial x_j^3 \partial x_k}(x_j, x_k) + \frac{h_k^2}{6} \frac{\partial^4 f}{\partial x_j \partial x_k^3}(x_j, x_k) \right] + \text{Higher order terms} \quad (4.28)$$

h for the Hessian equations is computed in subsection 4.2.7.

Comparing the first two methods, it can be seen that for h of the order of machine precision, they are almost equivalent. In the CONDUIT case, h is bigger because the performance accuracy is lower. A good choice for h will be seen in the next subsection. For a bigger h , the second method for computing the Hessian is more accurate than the first one. Both the truncation error and the roundoff error are smaller for the second case. The truncation error of method one includes the term $\frac{\partial^4 f}{\partial x_j^2 \partial x_k^2}(x_j, x_k)$. This term is usually bigger than $\frac{\partial^4 f}{\partial x_j^3 \partial x_k}(x_j, x_k)$ or $\frac{\partial^4 f}{\partial x_j \partial x_k^3}(x_j, x_k)$ because it contains lower-order partials in any parameter. In case that the third order partials are zero, this term may still exist. The strength of the second method consists in its symmetry. Both the first and the second method were coded and tested. The second method is recommended and used in the example presented in subsection 4.2.8.

4.2.7 Results on Accuracy of Hessian Computation

There are two sources of error in Hessian equations: truncation error and round-off error [18].

- **The roundoff error:** With h exactly representable, one has a roundoff error in method 2:

$$\epsilon_r \approx \frac{4\epsilon_f f}{4h^2} \quad (4.29)$$

where ϵ_f is the fractional accuracy with which f is computed.

- **The truncation error:**

$$\epsilon_t \approx \frac{h^2}{6} \left[\frac{\partial^4 f(x_j, x_k)}{\partial x_j \partial x_k^3} + \frac{\partial^4 f(x_j, x_k)}{\partial x_j^3 \partial x_k} \right] + O(h^4) \quad (4.30)$$

To minimize $\epsilon_r + \epsilon_t$, $\epsilon_r = \epsilon_t$. Therefore the optimal choice for h is

$$h = \left(\frac{6\epsilon_f f}{f''''} \right)^{1/4} = (\epsilon_f)^{1/4} x_c \quad (4.31)$$

where $x_c \approx (f/f''')^{1/4}$ is the “curvature scale” of the function f , or “characteristic scale” over which it changes. In the absence of any other information, one often assumes $x_c = x$ (except near $x = 0$).

In the CONDUIT case $h = (0.0002)^{(1/4)} = 0.1189$.

Remark 1. In order to have $\epsilon_f = 0.0002$ precision for specs they have to be read from the Pcomb chart with at least 4 significant digits instead of 2, and we made this modification in pcomb.c.

4.2.8 Example

An example of the case where the gradient of MAX_ACTIVE is zero is the optimal point found for the F-14 problem, starting at the trivial point and optimizing with $\epsilon_f = 0.0002$. The Pcomb chart shows that the optimal point is the local minimum of the spec aerrspc, which does not compete with any other spec, as can be seen in Figure 4.14:

The sensitivity analysis was done for this point, using the second method for the Hessian computation and step $h=0.1189$. The results found are displayed in Figures 4.15, 4.16, and 4.17.

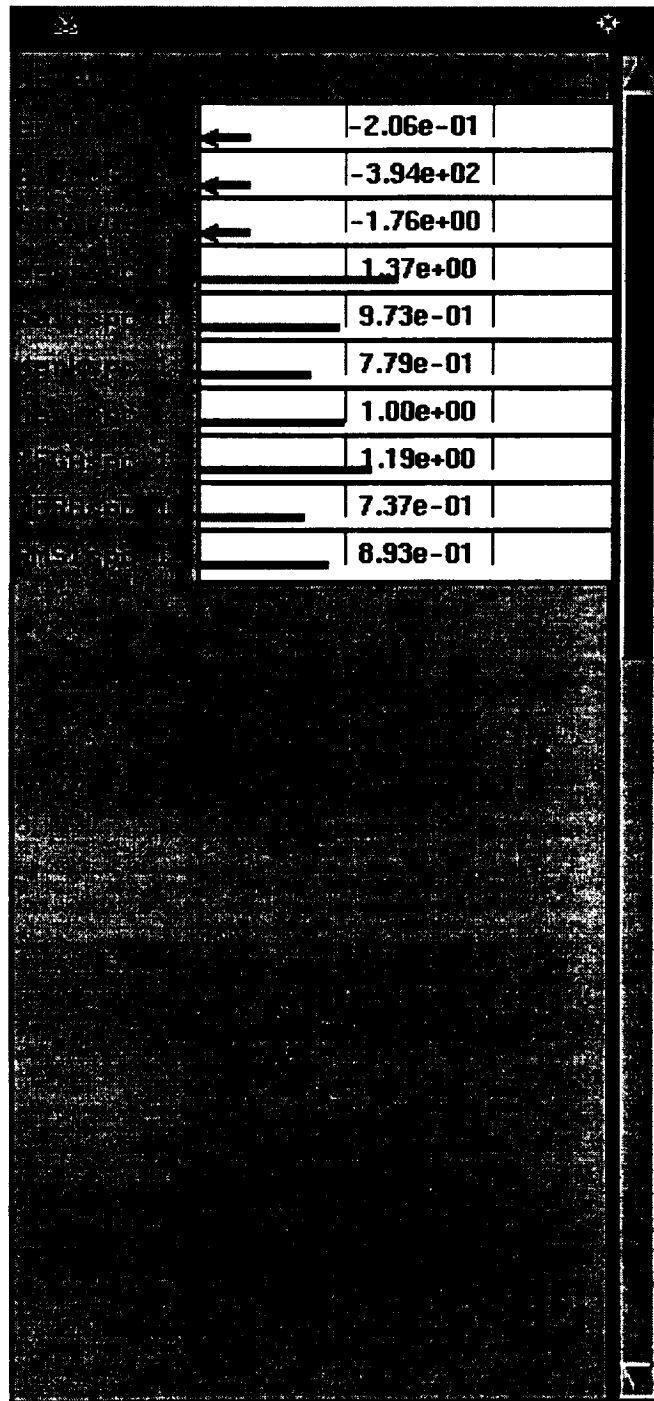


Figure 4.14: Grumman F-14, performance display in the Pcomb chart

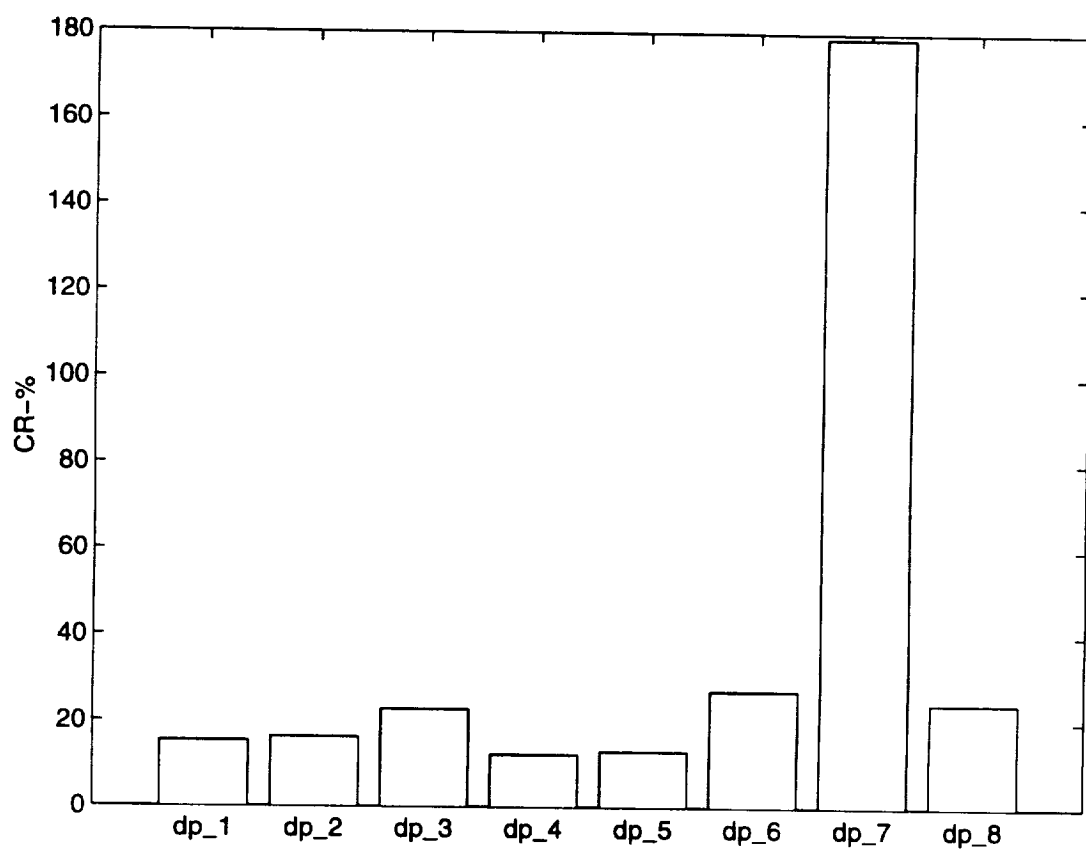


Figure 4.15: Grumman F-14, Cramér-Rao bounds

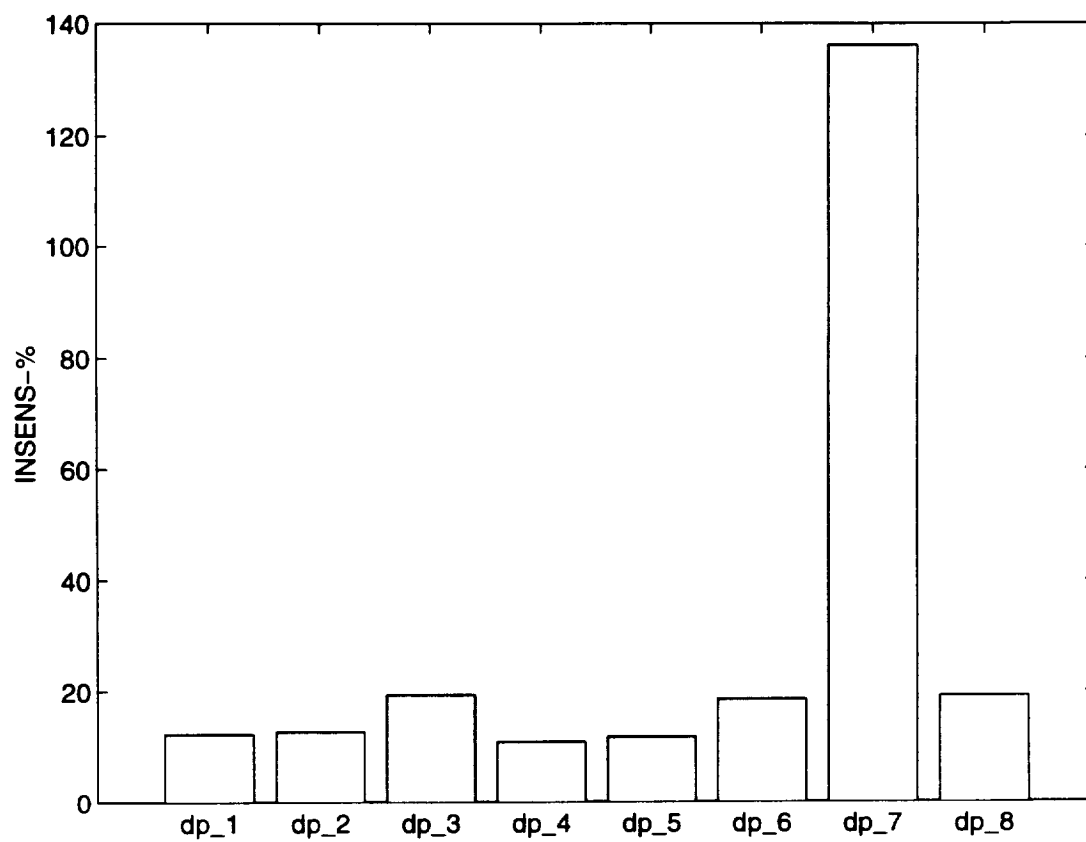


Figure 4.16: Grumman F-14, insensitivities

Confidence Ellipsoids for CR_PERCENT > 20%

	dp_a	dp_b	dp_7	dp_8
dp_a	-0.198155	0.168134	-0.126817	0.0316914
dp_b	0.233028	-0.197187	0.142168	-0.0491662
dp_7	1.17873	-0.227597	0.182727	0.43479
dp_8	0.078873	-0.221333	0.204325	0.0707156
dp_9	0.0343899	-0.0180513	0.147884	0.0635992
dp_6	-0.283666	1.46912	-0.918292	-0.678803
dp_7	0.203625	-0.82105	1.31355	0.357615
dp_8	0.468506	-0.586867	0.345797	1.27014

Figure 4.17: Grumman F-14, correlations

The largest Cramér-Rao bound corresponds to dp_7 and this is a consequence of a large insensitivity of the performance to this parameter. For the F-14 problem this does not necessarily mean that dp_7 is unimportant, but there is no spec to drive the parameter in this phase of the CONDUIT design run. dp_7 is the time constant in the stick prefilter and it is tuned to reject the noise added to the pilot command. Because in this simulation there is no noise added to the cockpit, performance is insensitive to dp_7. Sometimes a parameter appears insensitive because the solution is very robust in that parameter at the optimal point.

A much more reliable result is the correlation result. Figure 4.17 shows the correlation of design parameters having Cramér-Rao bounds larger than 20%. Every shown column corresponds to such a parameter and shows the correlations of this parameter with all the other design parameters. For this particular case dp_3, dp_6, dp_7, and dp_8 have Cramér-Rao bounds larger than 20% and they are analyzed for correlation. As was explained, dp_7 has Cramér-Rao bound large as a consequence of its insensitivity. But the others are sensitive parameters and their large Cramér-Rao bounds are a consequence of their correlations. Figure 4.17 shows that :

- dp_3 is correlated to dp_8
- dp_6 is correlated to dp_7 and dp_8
- dp_8 is correlated to dp_6, dp_3 and dp_7.

This suggests that the elimination of dp_8 may be a good approach. dp_8 is the integrator gain and dp_4 and dp_5 can do its job.

For the same optimal point we computed the eigenvalues and the eigenvectors of the scaled Hessian of the MAX_ACTIVE. The eigenvalues are shown

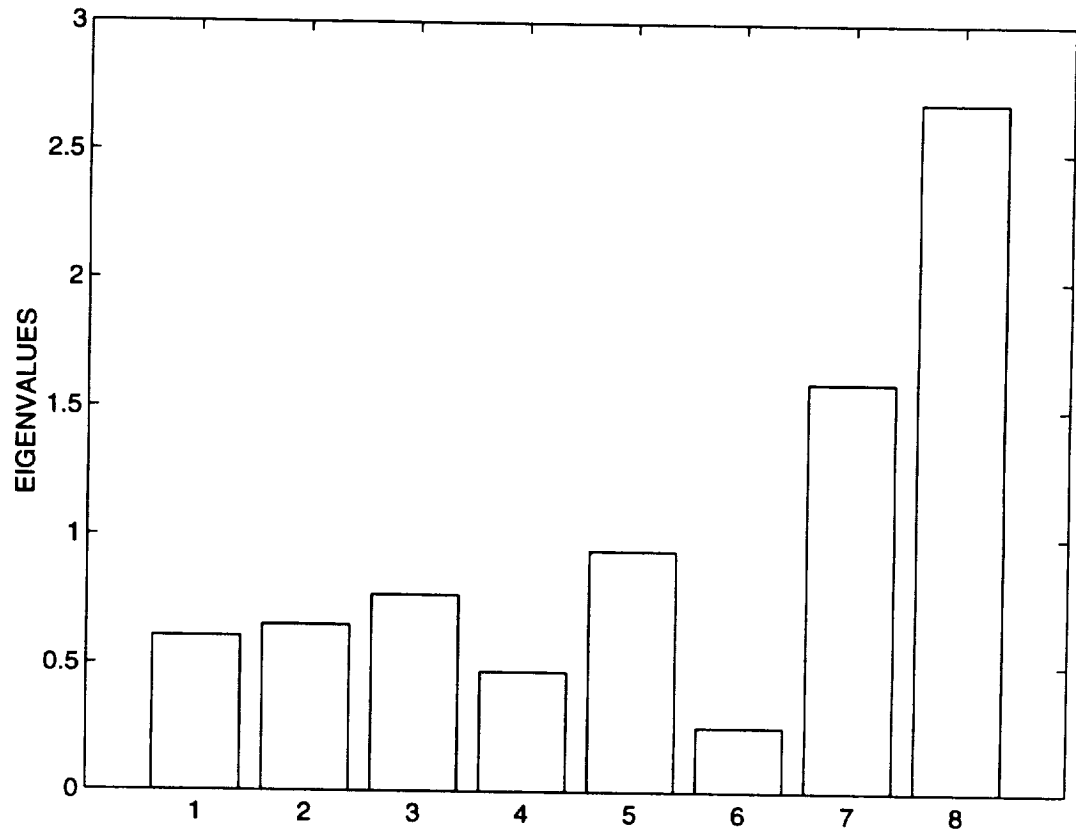


Figure 4.18: Grumman F-14, Hessian's eigenvalues in the optimal point

in Figure 4.18 The corresponding eigenvectors are displayed in Table 4.1. The eigenvalues plot shows that the largest eigenvalues are the seventh and eight, corresponding to dp_7 and dp_8. In the eigenvectors Table 4.1, every column is an eigenvector corresponding to a dp and shows the correlation of the dp with all the others. Note that dp_7 is highly correlated to dp_6 and that dp_8 is correlated to all dp's but dp_7. These results are in accordance with the results of the Cramér-Rao ellipsoid.

Table 4.1: Grumman F-14, Hessian's eigenvectors in the optimal point

1	2	3	4	5	6	7	8
vector	vector	vector	vector	vector	vector	vector	vector
0.5401	-0.2344	0.1045	0.6348	-0.1139	-0.0877	0.1197	-0.4522
-0.3502	0.1593	-0.1088	0.7643	0.1073	0.1112	-0.1182	0.4670
-0.4322	-0.5166	0.1063	0.0336	0.5737	0.2556	-0.0098	-0.3733
0.3676	0.0898	0.7463	-0.0468	0.3785	0.1420	-0.1344	0.3408
0.4760	-0.3160	-0.5830	-0.0851	0.3545	0.0379	-0.3376	0.2922
-0.0445	-0.2497	0.0209	0.0013	0.2438	-0.6692	0.5790	0.3045
0.1874	0.3423	-0.2363	-0.0159	0.2613	0.5206	0.6723	-0.0192
-0.0020	-0.6044	0.1132	-0.0454	-0.4997	0.4174	0.2289	0.3788

4.3 Gradient of MAX_ACTIVE is discontinuous at the optimal point

Consider the case of the constrained minimization problem

$$\begin{aligned}
 & \text{minimize } f(x) \\
 & \text{subject to } h(x) = 0 \\
 & \quad \quad \quad g(x) \leq 0
 \end{aligned} \tag{4.32}$$

$f : R^n \rightarrow R, h : R^n \rightarrow R^m, g : R^n \rightarrow R^p$, all C^1 , and the optimal point x^* . In case that there are more objectives, $f(x)$ is the maximum of them. If the optimization stopped in Phase 1 or Phase 2, when not all hard and soft constraints are satisfied, then the following discussion does not apply. It is true only for optimal points that satisfy all the constraints.

If the optimal point x^* is found in case B), the gradient of MAX_ACTIVE

is discontinuous, but the gradients of the objective and constraints satisfy the Kuhn-Tucker Conditions [21][22].

Definition 4.1 [21] *Let x^* be a point satisfying the constraints*

$$h(x^*) = 0, g(x^*) \leq 0 \quad (4.33)$$

and let J be the set of indices j for which $g_j(x^) = 0$. Then x^* is said to be a regular point of the constraints 4.33 if the gradient vectors $\nabla h_i(x^*)$, $\nabla g_j(x^*)$, $1 \leq i \leq m, j \in J$ are linearly independent.*

Theorem 4.2 (Kuhn-Tucker Conditions) [21] *Let x^* be a local minimizer for the problem 4.32 and suppose x^* is a regular point for the constraints. Then there is a vector $\lambda \in R_m$ and a vector $\mu \in R_p$ with $\mu \geq 0$ such that*

$$\nabla f(x^*) + \lambda \nabla h(x^*) + \mu \nabla g(x^*) = 0 \quad (4.34)$$

$$\mu g(x^*) = 0 \quad (4.35)$$

$$h(x^*) = 0 \quad (4.36)$$

The Kuhn-Tucker Conditions still hold when the regularity conditions for x^* are replaced by a weaker condition, i.e. the Kuhn-Tucker constraint qualification (KTCQ) satisfied at x^* [22]:

$$\overline{coTC(x^*, \Omega)} = \{h \mid \langle \nabla g_j(x^*), h \rangle \leq 0 \ \forall j \in J(x^*)\} (= S(x^*))$$

where the tangent cone is defined by:

$$TC(x, \Omega) = \{h \in R^n \mid \exists o(\cdot), \bar{t} > 0 \text{ s.t. } x + th + o(t) \in \Omega \\ \forall t \in (0, \bar{t}], \frac{o(t)}{t} \rightarrow 0 \text{ as } t \rightarrow 0, t > 0\}$$

and Ω is defined by Equation 4.35 and Equation 4.36, $co(X)$ is the union of all convex combination of finite subsets of X , and $S(x^*)$ is called the set of “first-order feasible “ direction [22].

By inspecting Equation 4.34 for the very simple and frequent case when at the optimal point the objective pushes one constraint to its border,

$$\nabla f(x^*) + \mu_j \nabla g_j(x^*) = 0, \quad j \in J(x^*)$$

it can be seen that the scalar product of the two normalized gradients is $\mu_j = -1$. In this case the other components of μ , corresponding to non-active constraints, are zero. Analyzing the normalized gradients of the objectives and constraints at the optimal point, the active constraints can be recognized.

In case that there are more active constraints at the optimal point, their gradients satisfy Equation 4.34, but it is harder to identify them among all the active and non-active gradients. When the analytical formulae for the objectives and constraints are known the solution of the system can be found by defining various combinations of active constraints, solving the system of $n+p+m$ equations with $n+p+m$ unknowns, and checking the signs of the resulting Lagrange multipliers. μ_j must be positive or zero.

Second-Order Conditions exist also for problem 4.32 at the optimal point.

Theorem 4.3 (Second-Order Necessary Conditions) [21] [22] *Suppose the functions $f, g, h \in C^2$ and that x^* is a regular point of the constraints 4.33. If x^* is a local minimizer for problem 4.32, then there is a $\lambda \in R^m$, $\mu \in R^p$, $\mu \geq 0$ such that Equations 4.34 and 4.35 hold and such that*

$$\langle t, \frac{\partial^2 L}{\partial x^2}(x^*, \lambda, \mu)t \rangle \geq 0 \quad (4.37)$$

on the tangent subspace of the active constraints at x^* , with

$$L(x, \lambda, \mu) = f(x) + \lambda h(x^*) + \mu g(x^*)$$

These conditions reveal the existence of a positive semidefinite Hessian on the tangent subspace:

$$\forall t \in \left\{ \frac{\partial h}{\partial x}(x^*) t = 0, \langle \nabla g_j(x^*), t \rangle = 0 \forall j \in J(x^*) \right\} \quad (4.38)$$

Testing if the Hessian of the Lagrangian L can be used instead of the Hessian of MAX_ACTIVE in computation of the Cramér-Rao bounds, insensitivities, and correlations, in a constrained optimal point, remains a suggestion for future work.

4.4 Parameter sensitivity to small changes in constraints

A sensitivity theorem exists to make sure that in the usual cases (functions are C^2) if an optimal solution satisfying all the constraints exists, then after small variations of the constraints an optimal solution continues to exist.

Theorem 4.4 [21] [22] *Let $f, g, h \in C^2$ and consider the family of problems*

$$\begin{aligned} & \text{minimize } f(x) \\ & \text{subject to } h(x) = c \\ & \quad \quad \quad g(x) \leq d \end{aligned} \quad (4.39)$$

Suppose that for $c=0, d=0$, there is a local minimizer x^ that is a regular point and that, together with the associated Lagrange multipliers, $\lambda, \mu \geq 0$, satisfies the*

second-order sufficiency conditions for a strict local minimum. Assume further that no active inequality constraint is degenerate. Then for every $(c, d) \in R^{m+p}$ in a region containing $(0, 0)$ there is a solution $x(c, d)$, depending continuously on (c, d) , such that $x(0, 0) = x^*$, and such that $x(c, d)$ is a local minimum point of 4.39. Furthermore,

$$\nabla_c f(x(0, 0)) = -\lambda \quad (4.40)$$

$$\nabla_d(x(0, 0)) = -\mu \quad (4.41)$$

Part IV

Conclusions and Suggestions for Future Work

In this report, two multivariable control system design problems were analyzed and optimized using CONDUIT. This section will summarize the conclusions that are drawn from the results presented in Chapter III and then suggest possible paths of continued research.

First we must emphasize that CONDUIT is an invaluable tool for aircraft control system design, that allows a quick analysis and design of different control problems. Every change in the design can be executed instantaneously because there is no need for code writing.

Although the two discussed problems are very different, they were satisfactorily optimized. The system in the F-14 problem is a rather slow system and was analyzed especially in the time domain, while the system in the antenna problem is fast and was analyzed only in the frequency domain.

Before starting the optimization process the models and the specifications had to be changed as follows:

- F-14 problem: the block diagram was improved by adding a wind gust model and the performance chart was supplemented with three specs - Bandwidth & Time delay, RMS g's at the pilot caused by turbulence, and the eigenvalues test
- Antenna problem: stability was insured by adding an eigenvalues test spec and stability margins spec, and by requiring positive design parameters.

The value of the discretization sample rate for the time simulation, the value of the gradient step-size, and different initialization points were tested. The sensitivity issues were addressed. The performance sensitivity to controller parameters changes was detailed and parameter insensitivity and correlation were

derived as an approximation of the estimation theory. These were illustrated for an optimal point achieved for the F-14 problem.

Some problems still remain open and should be addressed in future works:

- How can changes in the sample rate affect the optimal design and how sensitive is the optimal design to them ?
- Why does the optimal point vary slightly when starting the optimization in different points close to each other ?
- Is it possible to use the gradients or the Lagrangian instead of the Hessian of J in the constrained optimal point to compute Cramér-Rao bounds, insensitivities, and correlations of the parameters ?

Bibliography

- [1] Fan, M. K. H., Tits, A. L., Zhou, J., Wang, L.-S., Koninckx, J., *CONSOLE*, Technical Report, ISR, University of Maryland, College Park, 1990
- [2] Tischler, M. B., Colbourne, J. D., Morel, M. R., Biezad, D. J., Levine, W. S., Moldoveanu, V., *CONDUIT-A New Multidisciplinary Integration Environment for Flight Control Development*, AIAA Guidance, Navigation, and Control Conference, New Orleans, Louisiana, Aug. 11-13, 1997
- [3] Yudilevitch, G., Levine, W. S., *Techniques for Designing Rotorcraft Control Systems*, Technical Report, ISR, University of Maryland, College Park, 1994
- [4] Potter, P. J., *Parametrically Optimal Control for the UH-60A (Black Hawk) Rotorcraft in Forward Flight*, Master's Thesis, Department of EE, University of Maryland, College Park, 1995
- [5] Walrath, C. D., *Multivariable Optimization-Based Computer-Aided Design of Real Control Systems*, Master's Thesis, Department of EE, University of Maryland, College Park, 1983
- [6] Bishoff, D. E., *The Definition of Short-Period Flying Qualities Characteristics via Equivalent Systems*, Journal of Aircraft, V.20, No.6, June 1983, pp.494

- [7] Lin, C., Levine, W. S., Potter P. J., *GIFCORCODE User's Guide*, ISR, University of Maryland, College Park, 1995
- [8] Rimer, M., Frederick, D. K., *Solutions of the Grumman F-14 Benchmark Control Problem*, the Third IEEE Symposium on Computer-Aided Control System Design, Arlington, Virginia, September 24-26, 1986
- [9] Milne, G. W., *Grumman F-14 Benchmark Control Problem Solution using BLKLAB*, IEEE Control Systems Society Workshop on Computer-Aided Control System Design (CACSD), 1989
- [10] Anonymous, MATLAB 5.2., Simulink, Demos, Complex models, F-14 flight control simulation, Mathworks, Inc., 1998.
- [11] Soderstrom, T., Stoica, P., *System Identification*, Prentice Hall International (UK) Ltd, 1989
- [12] Franklin, F. G., Powell, J. D., and Workman, M. L., *Digital Control of Dynamic Systems*, Addison-Wesley Publishing Company, second edition, pp.455, 1990
- [13] Berman, H., and Gran, R., *Design Principles for Digital Autopilot Synthesis*, Journal of Aircraft, V.11, No.7, July 1974, pp.414-422
- [14] Stevens, L. B., Lewis, F. L., *Aircraft Control and Simulation*, John Wiley & Sons, Inc., pp.254, 1992
- [15] Lewis, F. L., *Applied Optimal Control & Estimation*, Prentice Hall and Texas Instruments, pp.249, 1992

- [16] Tischler, M. B., *Advances in Aircraft Flight Control*, Taylor & Francis Inc., pp.281, 1996
- [17] Maine, R. E., Iliff, K. W., *The Theory and Practice of Estimating the Accuracy of Dynamic Flight-Determined Coefficients*, NASA Reference Publication 1077, 1981
- [18] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. 1992, *Numerical Recipes in C*, (Cambridge Univ. Press), pp.188
- [19] Anon., *Aeronautical design standard, handling qualities requirements for military rotorcraft*, ADS-33C, August, 1989
- [20] Anon., *Military specification, flying qualities of piloted airplanes*, MIL-F-8785C, November, 1980
- [21] Luenberger, D. C., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, 1973
- [22] Tits, A. L., *Notes for ENEE 664: Optimal Control*, ISR, University of Maryland, College Park, 1996
- [23] Tischler, M.B. *Advancements in Frequency-Domain Methods for Rotorcraft System Identification*, Vertica 13, no. 3, pp. 327-349, 1989.
- [24] Morel, M.R., *CONDUIT Handling Qualities Specifications: Development, Unique Features, and Applications*, M.S. Thesis, California Polytechnic State University, 1979.

- [25] Tischler, M.D., Colbourne, J.D., Cheung, K.K., Frost, C.B., Levine, W.S., and Moldoveanu, V., CONDUIT “Control Designer’s Unified Interface” Course Notes, NASA/CP-1998-10157, USAAMCOM-TR-98-A-007, 1998.